

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 11:10:24 ; Search time 19 seconds
(without alignments)
1350.357 Million cell updates/sec

Title: US-09-744-226a-1
Perfect score: 4563
Sequence: 1 AEOTRNHLNAGDITYSVRAM.....KDGTPPEGSSKPAHLVTSLSL 872

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTIUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	927	20.3	240	US-09-370-098-5	Sequence 5, Appl1
2	776	17.0	652	US-08-956-322-2	Sequence 2, Appl1
3	775	17.0	835	US-09-284-819-6	Sequence 6, Appl1
4	774	17.0	652	US-09-110-116-1	Sequence 1, Appl1
5	769	16.9	521	US-08-956-322-4	Sequence 4, Appl1
6	715.5	15.7	886	US-09-110-116-3	Sequence 3, Appl1
7	652.5	14.3	344	US-09-110-116-4	Sequence 4, Appl1
8	521.5	11.4	1052	US-08-852-806-2	Sequence 2, Appl1
9	521.5	11.4	1052	US-09-163-669-2	Sequence 4, Appl1
10	517	11.3	231	US-09-370-098-4	Sequence 2, Appl1
11	497	10.9	235	US-09-370-098-6	Sequence 6, Appl1
12	475	10.4	884	US-08-465-976A-2	Sequence 2, Appl1
13	475	10.4	884	US-08-982-412-2	Sequence 2, Appl1
14	290.5	6.4	415	US-08-110-286A-6	Sequence 6, Appl1
15	290.5	6.4	415	US-08-981-189B-10	Sequence 10, Appl1
16	290.5	6.4	415	US-08-482-746-15	Sequence 6, Appl1
17	290.5	6.4	415	US-08-482-746-13	Sequence 6, Appl1
18	289.5	6.3	415	US-08-110-286A-2	Sequence 2, Appl1
19	289.5	6.3	415	US-08-482-746-2	Sequence 2, Appl1
20	279.5	6.1	228	US-09-724-864-42	Sequence 42, Appl1
21	273	6.0	444	US-08-482-746-15	Sequence 15, Appl1
22	261.5	5.7	431	US-08-381-433A-2	Sequence 11, Appl1
23	261.5	5.7	431	US-08-981-189B-11	Sequence 11, Appl1
24	261.5	5.7	431	US-08-482-746-10	Sequence 10, Appl1
25	260.5	5.7	411	US-08-381-433A-4	Sequence 4, Appl1
26	260.5	5.7	411	US-08-981-189B-12	Sequence 12, Appl1
27	260.5	5.7	431	US-08-981-189B-13	Sequence 13, Appl1

28	259.5	5.7	462	US-09-238-796-2	Sequence 2, Appl1
29	245.5	5.4	449	US-09-414-189-1	Sequence 1, Appl1
30	244.5	5.4	425	US-08-795-876-44	Sequence 44, Appl1
31	241.5	5.3	448	US-08-811-897A-22	Sequence 22, Appl1
32	241.5	5.3	448	US-08-855-213-22	Sequence 22, Appl1
33	241.5	5.3	448	US-09-201-474-22	Sequence 22, Appl1
34	241.5	5.3	525	US-08-811-897A-23	Sequence 23, Appl1
35	241.5	5.3	525	US-08-855-213-23	Sequence 23, Appl1
36	241.5	5.3	525	US-09-201-474-23	Sequence 23, Appl1
37	240.5	5.3	476	US-08-811-897A-24	Sequence 24, Appl1
38	240.5	5.3	476	US-08-855-213-24	Sequence 24, Appl1
39	240.5	5.3	476	US-09-201-474-24	Sequence 24, Appl1
40	240.5	5.3	553	US-08-811-897A-25	Sequence 25, Appl1
41	240.5	5.3	553	US-08-855-213-25	Sequence 25, Appl1
42	240.5	5.3	553	US-09-201-474-25	Sequence 25, Appl1
43	240	5.3	474	US-08-453-742-2	Sequence 2, Appl1
44	240	5.3	474	US-08-454-464-2	Sequence 2, Appl1
45	240	5.3	474	US-08-453-222-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-370-098-5
; Sequence 5, Application US/09370098
; Patent No. 6303768
; GENERAL INFORMATION:
; APPLICANT: Lin, Yi-Yun
; APPLICANT: Benzer, Seymour
; APPLICANT: California Institute of Technology
; TITLE OF INVENTION: METHUSELAH GENE, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 06618/343001
; CURRENT APPLICATION NUMBER: US/09/370,098
; CURRENT FILING DATE: 1999-08-06
; PRIOR APPLICATION NUMBER: 60/095,826
; PRIOR FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-370-098-5

Query Match      20.3%; Score 927; DB 4; Length 240;
Best Local Similarity 67.5%; Pred. No. 6.4e-76;
Matches 162; Conservative 46; Mismatches 32; Indels 0; Gaps 0;

QY      285 HDLLDYTWVGLSLVLCILICIFCFEFGIQSQRNTIHKNLCLSLFVAELFLIGIN 344
      ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
DB      1 NEILLSVITWVGIVISLVCLALCISFCEFLRGIDTRNTIHKNLCLINLFAELFLVIGID 60

QY      345 RDPQACAVFALLHEFFLAFTWMEELGVOLYIMLVEFESEHSRRRYFYLVGMPA 404
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      61 KQYEVACIFPAGLHYFLFLAFAFSWCLDGVHLYLLVEFESEYSRTYYIIGGCFPA 120

QY      405 LIVAASAAVDYRSGTDKWCWRLDTYFTWSPFGATLLIMLVIFGLALYKMFHTAI 464
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      121 LVVGIAAIDYRSGTEKACWMLRVNDYFTWSPFGATLLIMLVIFGLALYKMFHTAI 180

QY      465 LKPESCDINIKSWVGAIALCLGLTMAFGMTINESTVIMAYLFTFNISQGFITI 524
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      181 LKPDSSRLDNIKSWALGAILLFLGLTMAFGMLFTFNKSSVWAYLFTFNMAQGFIV 240

RESULT 2
US-08-956-322-2
; Sequence 2, Application US/08956322
; Patent No. 6277977
; GENERAL INFORMATION:
; APPLICANT: SATHE, GANESH
```

```

APPLICANT: MAO, JOYCE
TITLE OF INVENTION: CDNA CLONE HAP0167 THAT ENCODES
TITLE OF INVENTION: A HUMAN 7-TRANSMEMBRANE RECEPTOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,322
FILING DATE: 23-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/049,329
FILING DATE: 11-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70075
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-956-322-2

```

Query Match 17.0%; Score 776; DB 4; Length 652;

Best Local Similarity 33.6%; Pred. No. 1.8e-61;

Matches 184; Conservative 102; Mismatches 229; Indels 32; Gaps 11;

```

QY 35 TPGKDSAASTSLNKAMETVNNLLQPOLNAMRDLTSDQLRAATMLLHYEESAFVLAD 94
DB 120 TTSSKTQGGKELQKIDKRESLLTNTLT--WRREGQELSSSTATTILRDVESVLEFAL 177
QY 95 NLKTDIVRENTDNKLEVARLSTEGNLEDKPEPNMGHSTIQLSANTLKQNGRGEIR 154
DB 178 KDEQKVLKIDNSVAIETQITDNCSEERKTFNLVQNNMS-MDIRCSDIQDTPGPSA 236
QY 155 VAFVLVNLGPRYLSTENASKKLGTLEALSTHNSVIVNSPVYTAATNFKSNKYIADYVF 214
DB 237 IATLSTSSLSLNIIT--NATP--FEEMDKDQVYLNQVVAALGPR--RNVSLSKSVTL 288
QY 215 TVRHKQSEENFNPNCSFMSYSKRTMTG-YWSTQGCRLTNTKHTTSCNHLNFAVLM 273
DB 289 TFGHVAKTPTETKAYECYWM--KSTGGSGMSRGCFLIHVNSHTKNCNCHLSFPAVL 345
QY 274 AHVEVKHSDAVHDLDDVITWVGILSLVCLLICTFCEFRGLQSDRNTIHKMLCISLF 333
DB 346 ALTSQE-----EDPVLVITVYVGLSVSLCLLALFLCKAIONSTSLHQLDLSCLF 400
QY 334 VALLFLIGINTDQIPACAVFAALLHFFFLAFTWFLGVOLYT---MLVEVFSEH 389
DB 401 LAHLFLVGLDREPKVLCIIAGALHYLYLAFTWMLLGGVHLFTLARNLTVVNSIN 460
QY 390 SRRRY-FYLVGYGMPALIVSAVADYRGTQVCWLRDLTYEIMSGPATILIMLVN 448
DB 461 RLKWKIMEPVGYPAVTVAISAASPHLYGTDRCLWHLDDGGHMSFLGVCVAIFSA 520
QY 449 IFGLIALYKMFHHTALIKPESGCLDNKISWVIGALIALLLGLTWAGLMIINESTV 508

```

```

DB 521 VLFVIFVWLKRLKLSINSEVSTIONTRMLAFKATAQLFLLIGTCWGLQVGPAAVMA 580
QY 509 YLFTFNSLQGMFFTHCVLQKKVREYKGLCLTHCCSGKSTSSIGSGK-----TS 561
DB 581 YLFTFINSLOGFFLEFLVYCLLSQVOKOYOKWFEIRVKSSESEYTLTSSKMGPDSP 640
QY 562 GSRTPRG 568
DB 641 GDVPRQG 647

```

RESULT 3

US-09-284-819-6
Sequence 6, Application US/09284819
Patent No. 6365712

```

GENERAL INFORMATION:
APPLICANT: Kelly, Kathleen
APPLICANT: The Government of the United States of America
APPLICANT: as represented by The Secretary of the
APPLICANT: Department of Health and Human Services
TITLE OF INVENTION: Methods and Compositions for Inhibiting Inflammation
TITLE OF INVENTION: and Angiogenesis Comprising a Mammalian CD97 Alpha
FILE REFERENCE: Subunit
CURRENT APPLICATION NUMBER: US/09/284, 819
CURRENT FILING DATE: 1999-08-20
EARLIER APPLICATION NUMBER: US 60/027, 871
EARLIER FILING DATE: 1996-10-25
EARLIER APPLICATION NUMBER: WO PCT/US97/19772
EARLIER FILING DATE: 1997-10-24
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 835
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: CD97amino acid sequence encoded by full-length
US-09-284-819-6

```

Query Match 17.0%; Score 775; DB 4; Length 835;

Best Local Similarity 33.3%; Pred. No. 3.4e-61;

Matches 191; Conservative 108; Mismatches 211; Indels 64; Gaps 18;

```

QY 30 QLRNLTPGCGDSAASTSLNKAMETVNNLLQPOLNAMRDLTSDQLRAATMLLHYEES 88
DB 281 KYVDLGRDSKSTSAEVTIQVYIKLVDELMEAPGDVEA--LAPVRHLITQLLSNLEDI 337
QY 89 AFVLADNLKTDIVRENTDNKLEVARLSTEGNLEDKPEPNMGH-STIQLSANTLKON 147
DB 338 MRIAKSLPRGPFYIYSPSTELTL-MIOERGD-----KNYMGSSAKMLINAVAA 389
QY 148 GRN-GEIRVAFVLYNNGPYLSTENASKKLGTLEALSTHNSVIVNSPVY-----TAAT 198
DB 390 GAEDPGAVAGIISIQWMTLLA--NASLNIHSKQKQELBEIYESS--INGVOLRLSLAV 445
QY 199 NKEP---SNKYIADPVYFVFKHKQSE-----ENENPN-----CSFMSYSKRTM 240
DB 446 NSIFLSHNNTKEINSPILFAFSLHSESDGAGRDPAPKADWPGPRQELLCAFMK-SDSDR 504
QY 241 TGWSTQGCRLTNTKHTTSCNHLNFAVLAHVEVKHSDAVHDLDDVITWVGILSL 300
DB 505 GGHMATEGCOVLGSKNGSTTCQCSHLSFALIMAHYVE-----DKKLTLIRVGLALS 538
QY 301 LVCLLICTFCEFRGLQSDRNTIHKMLCISLFAELLFLIGINTDQPIA--CAVFAAL 358
DB 559 LFCLLCITLFLVLRPIQSGRTIHLHLCLCFVGSITFLAGIENEGGVGLCRLVAGL 618
QY 359 LHFFFLAFTWMLLEGVOLYIMLVEVFESHSRRKRYFLVGYGMPALIVSAVADYRST 418
DB 619 LHYCFLAAPCMMSLEGEFLVAVRVGOGGLSTRMLCLIGYGVPLIVGSAIYSKGY 678

```

[illegible]

```

RESULT 4
US-09-110-116-1
; Sequence 1, Application US/09110116
; Patent No. 6013479
; GENERAL INFORMATION:
; APPLICANT: Xu, Hong
; APPLICANT: Cohan, Victoria L.
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: HUMAN EKR1-LIKE G PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR
; FILE REFERENCE: PF-0550 US
; CURRENT APPLICATION NUMBER: US/09/110,116
; CURRENT FILING DATE: 1998-07-02
; NUMBER OF SEQ. ID NOS.: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 652
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: 429905, EOSINOT03
; US-09-110-116-1

```

	Query Match	17.0%	Score 774:	DB 3:	Length 652:
	Best Local Similarity	33.6%	Pred. No. 2.7e-61:		
	Matches 184:	Conservative 102:	Mismatches 229:	Indels 32:	Gaps 11:
QY	35	TPGGKDSAAKSLINKAMWETVNNLLQPOLANMRDLTTSDDQRAATMLHTVEESAFVLAD	94		
DB	120	TTSSKTTQGRKELQIKYDKFESLLTNGTL-WRTGREGIESTATITLIDRESKLYETAL	177		
QY	95	NLAKTIDVIREMTDNKLEVARLSTEGNLEDLKFEPMNHGSIQASNTLQNGNGEIR	154		
DB	178	KDPEKQVNLKIONDSVALETQAITDNCSEERKTFNLVNOYMS-MDRICSDIITGGDTQGPSV	236		
QY	155	VAFVLYNNLGPYILSTENASMKLGTALSTNSHVIYNSPVIATAIKKEFSNKKYLLADPVVF	214		
DB	237	IATFSYSSLGNI--NATF--FEEMDKQDQVYVYNSQVSAIGPK--RNVSLSKSVTL	288		
QY	215	TYKHIIKQSEENFNPNCSFMSYSKRPMTNG-VASTQGCRLTTNKHTYTTSQCNHLTFAYILM	273		
DB	289	TFQIAHKMPTSTKKYPCVYW---KSTGGQSQMSRDCFLIHVKKSHTCMCSHLSEFAVILM	345		
QY	274	ARHEVKHSDAVIDLLDVIYTWGILLSVYCLLICIFTECFPRGLQSDRNTIHKNCISLDF	333		
DB	346	ALTSOE-----EDPLVITVYGLSVSLCLLLALFTLCKAIQNTSTSHLDQSLCIE	400		
QY	334	VAEILLFLGINRTQPIACAFNALHREFFLAAFTWMLBEVQYI----MLVEYFESEB	389		
DB	401	LAHLLEFLVIGIDRTEKVCYLCSTIAGLAHLVLYLAFTWMLBEVHLEFLTARNLTIVNYSIN	460		
QY	390	SRRRY-FYLVGYGMPALIVAAVAANDVRSYSTDGKCMRLDTFYFMSEIFPATLITMLNV	448		
DB	461	RLMKWIMPEVGYGDAVYVVAISAASWPHLYGADRCWCHLHDGGEFMSFLGPYCAIFSANL	520		
QY	449	IFELGALYKMHNRHAIKLPESGCIDNKSWSYIGAILALLCLIGLTAFGLMTINESTVIMA	508		
DB	521	VFLIVLEWILIRKLSSLNSEVSTIONTMYLAFKAAVQALFELIGCTWICLILDOVGPAAQVMA	580		

OY	509 YLFIIFPISLOGMFEIHFHCVLOKREKYGCGLTRHCSCSGSTSSIGSK-----TS	561
	: : : : :	
Dd	581 YLFIIINSLSGFFFLFVAYCLLSQGVQNOYOQMRFNEIYKSSESETYTLSSKMGPDKPSE	640
OY	562 GSRTPRGR	568
	:	
Dd	641 GDVEPGQ	647

```

: RESULT 5
: US-08-956-322-4
: Sequence 4, Application US/08956322
: Patent No. 627977
: GENERAL INFORMATION:
: APPLICANT: SATHE, GANESH
: APPLICANT: MAO, JOYCE
: TITLE OF INVENTION: CDNA CLONE HAP0167 THAT ENCODES
: TITLE OF INVENTION: A HUMAN 7-TRANSMEMBRANE RECEPTOR
: .NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: RATNER & PRESTIA
: STREET: P.O. BOX 980
: CITY: VALLEY FORGE
: STATE: PA
: COUNTRY: USA
: ZIP: 19462
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/956,322
: FILING DATE: 23-OCT-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/049,329
: FILING DATE: 11-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: PRESTIA, PAUL F
: REGISTRATION NUMBER: 23,031
: REFERENCE/DOCKET NUMBER: GH-70075
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-407-0700
: TELEFAX: 610-407-0701
: TELEX: 846169
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 521 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-956-322-4

```

[illegible]


```
/ Patent No. 5874245
/ GENERAL INFORMATION:
/ APPLICANT: Shoji Fukusumi
/ APPLICANT: Shoji Hinuma
/ TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: SmithKline Beecham Corporation
/ STREET: 709 Swedeland Road
/ CITY: King of Prussia
/ STATE: PA
/ COUNTRY: U.S.A.
/ ZIP: 19406-0939
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/852,806
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/017,915
/ FILING DATE: 16 MAY 1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Han, William T
/ REGISTRATION NUMBER: 34,344
/ REFERENCE/DOCKET NUMBER: TAK50002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 610-270-5219
/ TELEFAX: 610-270-5090
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1052 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ US-08-852-806-2

Query Match 11.4%; Score 521.5; DB 2; Length 1052;
Best Local Similarity 22.3%; Pred. No. 5,2e-38;
Matches 233; Conservative 142; Mismatches 346; Indels 323; Gaps 43;

QY 10 AGDITYSVAMQVLGLVGLVLRNL-----PGKDSARSLNKAMETVNNLDP 60
DB 133 SGLDLFV-----DI-LRNVDTFKRAIYPSADV-----QREFFQVSEWDA 175
QY 61 QALNARLDITTSQDLRAATM-LHIVE-----ESAFVLADNLL----- 97
DB 176 EKKERKMD---AQVSPGSHVHLRYVEDLHLVGDAKAFQSSLIYTDNLVLSIOREPV 232
QY 98 -----KTDIVRENTDNKL-EVARLSTEGNEDL-----KF 127
DB 233 AVSSDITFPMRGRRGKMDVVRHSEDRFLPKFVLSLSPGKPAISGAAGSPGRGPGIV 292
QY 128 PENMGHSTIOLSANLTKONGRGEIRVAFVLYLNIGPLSTENASMKLGTEALSTNHXY 187
DB 293 PPGSGHSHORLLPADP-----DESSYFVIGAVLYRTLGLLPPRPPL----- 335
QY 188 IYNSPVTAAINKFEENKRYLADPVVETVKHIKOSEENFNPNCSFYSYKR-TMGWGS 246
DB 336 ATYSRMTYTVRPPROP---AEPIL-YVELSTIINGTDPKASADYSRADASSGDMPT 391
QY 247 QGCRLLTINKTKHTTCSNHLTNFAVILMAHVEVKSDAVDLLDVTWGIILSLVCLLI 306
DB 392 ENCGTLETOAHTRCOCCHLSTFAVL-----AQPRLDTLELAGSPSVPLVICAVS 443
```

```
QY 307 CIEFTF-----CFERGIOSDRNTIHKNICISLFVAELLFLIGINRTDQPIACVFAILL 359
DB 444 CMALLTLTAIYAFWRFISERSIILLNCLSLTILANLILVQSHVLSKGVCTMTAAFL 503
QY 360 HFFFLAFTMPLEGGVQLYIMLVEFESEHSRRKRYLYVGYCPALIVASAA-VVYRSY 418
DB 504 HFFFLSPCWVLEAMQSLAVIGRMTRLYRRR-FLCLQWGLPALVAVSVGFTTKCY 562
QY 419 GTDKVCLRLDYFIWSTIGPATLITMLNVIELGIALYKMFHHTALIKPESGCLDNKSM 478
DB 563 GTSYCYMLSEGLLYAFVGPAAVIVLNNML-IGIIVFNK-----LMARDGISDKSKQ 615
QY 479 VIGA-----IALLCIGLWARGMLYINE-STYIMAYLFIINSLOCMFIIFHYCLK 532
DB 616 RAGASLWSSCVPLPLALTMTMSAVLMTDRSVLFOALRAVFNAGVITAVHCFRLRE 675
QY 533 VRKEYGCLRTHCSCGSTESSIGSGKTSRTPGRYSGQSQRIRRMNDTVRKQSESS 592
DB 676 V-QDVYKCMQVGRADESDE-----PDSCKNGQLIILDFEKDV----- 714
QY 593 FITGDINSSASLNRBGLNARDTSVMDLP-----LNGNHN-SYSTASGE 638
DB 715 ---DIACOTVLFKE--VWTCNPSTYTGLSRSLDEDEPKSCLVGPBGSLSFPLPGN 768
QY 639 YL-----SNCVQIIDRCYNHETALEKKILKELTSNITPSYLNHNS 681
DB 769 ILVPMASPOLGEPPPPQENAPVYKCGEG-----LNQIDL---TWLRPTERG 813
QY 682 SEQNRMIMNKLYVNNL-----GSGREDDATVLDAT-----SFNHEESLGL 721
DB 814 SEGDVNVLPRLTSLDPPGGGGGGEDAPRARPGCTRRRAKTYAHTEGYPSFLSVHSG 873
QY 722 EL-----IHESDAPLLPRV-----Y 738
DB 874 GLGPAYGSLQNPYGMTFOPPPPTPSARQVDEPERGSRTPRGSTMRKGLERKRLY 933
QY 739 STENHQPHNYTRRIRIQQDSESFFPLTNEHTDLOSPHRDLSYTMPL-----A 789
DB 934 SLDLFKVMYTRKR--HSELVHELNOKFRTFD-----RYRSQSTAKRKRKSVSS 981
QY 790 GVATGESVTSQTPEPPPAKCGDAEDVYKSMPLGSRNHHQLHLYYOLRGSSDGFY 849
DB 982 GGAERSVCTD---KSPSE-----RSLSQHRHQSSTFKSMITLGS-----L 1022
QY 850 PPN-----KDGTPPEG 860
DB 1023 PKPRERLTLHRAAMEPTEPPDG 1046

RESULT 9
US-09-163-669-2
/ Sequence 2, Application US/09163669
/ Patent No. 6111076
/ GENERAL INFORMATION:
/ APPLICANT: FUKUSUMI, SHOJI
/ APPLICANT: HINUMA, SHOJI
/ TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED
/ TITLE OF INVENTION: RECEPTOR (HIBD07)
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Ratner & Prestia
/ STREET: P.O. Box 980
/ CITY: Valley Forge
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19482
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
```



```

RESULT 11
US-09-370-098-6
Sequence 6, Application US/09370098
Patent No. 6303768
GENERAL INFORMATION:
APPLICANT: Lin, Yi-Jyun
APPLICANT: Benzer, Seymour
APPLICANT: California Institute of Technology
TITLE OF INVENTION: METHUSELAH GENE, COMPOSITIONS AND
METHODS OF INVENTION: METHODS OF USE
FILE REFERENCE: 06618/343001
CURRENT APPLICATION NUMBER: US/09/370,098
PRIOR FILING DATE: 1999-08-06
PRIOR APPLICATION NUMBER: 60/095,826
PRIOR FILING DATE: 1998-08-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 235
TYPE: PRT
ORGANISM: Mus musculus
US-09-370-098-6

Query Match          10.9%; Score 497; DB 4; Length 235;
Best Local Similarity 40.9%; Pred. No. 6,6e-37;
Matches 96; Conservative 43; Mismatches 92; Indels 4; Gaps 1;

QY 298 LLSLVCLLCIFTCFPGFNGLOSDRNTIHKNLCLSLFVALLFLIGINTDPIACAVFAA 357
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 1 VSLVCLALAIATFTLLCRAVQNHNTYMHLLCVCFLAKILFLTGIDKIDNQTACALTAG 60

QY 358 LHFEEFLAFTMPLEGOVLIYIM----LVVEPSEHSRRKRYLYLVGYGPAALIVASAAV 413
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 61 FLHYFLCLCFMVLAEVWLEFVNRNLKVYVYSSRNKMLHLCAFGYGLPLVLIISASV 120

QY 414 DYRSYGTQVLCWLRLDTYFVMSFGPATLILMLNYFLGIALYKMHHTAILKPESGCLD 473
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 121 QPRGGMNRCWLNTETGTWSEFLGVCMLITIMSVLLAMTWLVRKILCSYSSEYSKIK 180

QY 474 NIKSVGAIALLCLGLTWAGLMYINVESTIVIMAYLFTIFNSLOGMFIIFHCY 528
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 181 DTRLTFRAIAQIFILGCSMWLGIFQIGPLASIMAYLFTIINSLOGAFIFLIHCL 235

RESULT 12
US-08-465-976a-2
Sequence 2, Application US/08465976a
Patent No. 5869632
GENERAL INFORMATION:
APPLICANT: SOPPET, DANIEL R
APPLICANT: LI, YI
APPLICANT: ROSEN, CRAIG A
APPLICANT: RUBEN, STEVEN M
TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN GILFILLAN, CECCHI
ADDRESSEE: STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NJ
COUNTRY: US
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,976a
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424

```

```

ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY P
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 994-1700
TELEFAX: (201) 994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 884 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-976a-2

Query Match          10.4%; Score 475; DB 2; Length 884;
Best Local Similarity 24.2%; Pred. No. 6,4e-34;
Matches 193; Conservative 106; Mismatches 306; Indels 192; Gaps 27;

QY 189 VNSPYITAAI--NKEFSNKVYLADPVFTVKHIKQSEENFNCSFWS-YSKRTMTGYWS 245
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 1 MNSPVSAVAVFGRNFFLNGI-LESPISEER-LIDPTANRKAICVQWDPDGLAEQHGWT 58

QY 246 TQGRRLTNKTKHTTQSCNHLTNFAVLMANVEKHSDAVHDL-LDVLITWVGLLSVCL 304
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 59 ANDCELVHNHNSHARCRCRGTGTVLMDASPRENLEG--DELLAVFTHHVAVANVAAL 116

QY 305 LLICTFCFPGFNGLOSDRNTIHKNLCLSLFVALLFLIGINTDPIACAVFALLHFEFL 364
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 117 VLTAAILLSTRLSKSNVRGIHANVAALGVALLLEFLGIGHRTNOLDVCTAAVALILLHYEFL 176

QY 365 AAFYTMFLEGOVLIYIMLVVEPSEHSRRKRYLYLVGYGPAALIVASAAVDYSGTDKVC 424
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 177 STEFAMLFVQGLHLYRQVEPRNVDRGAMREYHALMGVAVLLGLAVGIDPREGYGNPDFC 236

QY 425 WRLDTYFMSFGPATLILMLNYFLGIALYKMHHTAILKPESGCLDNIKSWYI--G 481
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 237 WISVHEPLWSFAGVPLVLIYVNN-----GTMLFARITSCSGQRAKKTSLTLRS 288

QY 482 AIALCLGLTWAGLMYINVESTIVIMAYLFTIFNSLOGMFIIFHCYLQKRYKEY-GKC 540
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 289 SFLLILVSAWSLFGILAVNHSILAFHYHAGLCGLGLAVLLPCVLNADARAAMPAC 348

QY 541 LFTHCCS-----GKSTSSIGSGKTSGRTPGRYS 570
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 349 LGRKAPEEPARPAIGPGAYNNNTALFEESGLIRITLGASTVSSVSAHS--GRTOODOS 406

QY 571 TGSQRIRRMNDTVRKQSESEF-----ITGDINSSASLNR 606
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 407 QGRGSLYLRD--NVIVRHGSAADHTDHSLOAHAGPTDLDVDMFHRDAGADSDSDLSLEE 464

QY 607 EGLLN-----NARDTSVMDTLR--LNGNHGNSYSIASGE-YLSN 642
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 465 ERLSLIPSESEDNRTGRGRFORPLCRAGQSRLTHPDVGDNDLSYWPALGDEEAP 524

QY 643 C-----VQIIDGVNHN-----ETAL-----EKKILK-ELTNSYIPS 673
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 525 CALQTWGSEERRIGLDTSKDANNNDPDPALTSGEDFSLGSAORGRKGIKNRLQYPLVQ 584

QY 674 YLNHRSSEQRNLMLNKLNNILGSGR---BDALVIDDASFVHNEESLGLHILH--- 725
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 585 TKGAPELSMCRATYLGHRVAPPAASYGRITAGGCTGSLSPARYSRSREDLLLRQLRS 644

QY 726 ---EESDAPLPP--RVYSTE-----NHQPHYTRRRIFQDHSSEFFPLLTJNEH 769
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 645 ERLERAPAVLPLRLSPGSOECMDAAPGRLEPKDKGSTLPRQPPRDYGCAGAFRGSHD 704

QY 770 TTDLOSPHRDSLYTGMPTLAGVAATESYTTSTQTEPPPAKCDADADVKKSPNIGSRNH 829
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
DB 705 ALDLCAP-REWL-----STLPPPRTRDLD----- 728
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 830 VHQLTHTYYQLGGRSSDG 846
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

Db 729 -HSPHLCPLPSGNSG 744

RESULT 13

US-08-982-412-2

Sequence 2, Application US/08982412
Patent No. 5938729

GENERAL INFORMATION:

APPLICANT: SOPEET, DANIEL R

APPLICANT: LI, YI

APPLICANT: ROSEN, CRAIG A

TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

STREET: 9410 KEY WEST AVENUE

CITY: ROCKVILLE,

STATE: MD

COUNTRY: US

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/982,412

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: BROOKES, ANDERS A

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PF181PCT2

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 884 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-982-412-2

Query Match 10.4%; Score 475; DB 2; length 884;

Best local Similarity 24.2%; Pred. No. 6,4e-34;

Matches 193; Conservative 106; Mismatches 306; Indels 192; Gaps 27;

Db 189 VNSPVITAI-NKESNKKYADPVVTVKHKIPOSENPNPCSFWS-YSKRTMTGWS 245

Db 1 MNSPVVSAVAFHGRNFRGI-LESPISLEFR-LIQTNRSKATCYQMDPGLAQBHGVT 58

Db 246 TGGCRLLTNTKHTTSCNLTNFAVMAVEKHSDAVDL-LDYITVWGIILSVCL 304

Db 59 ARCELVHRNGSHARCSCSTGTFGVLMDSPPRELEG-DLELLAVTHVVAVSAAL 116

Db 305 LICIFTECFRGLOSDNTHKNCISLFAVELLFLGINRTDOPICAVFAALLHFFFL 364

Db 117 VLTAAIILSLRSLSKSNRGHANVAALGYABELLFLGIRTNQVCTAVAILHFFFL 176

Db 365 AAFTMFLEGVOLYIMLEVEFESEHSRKKYFYLGVGMPALLIYAVSAVDYRSYTDKVC 424

Db 177 STEFAMLEVGQILHYRMQVBERNDRCAMRFYHALGMCVPAVLLGLAVGLDPEGVNDPFC 236

Db 425 WMLDPTFTSFITGPALITLMLNVIPIGIALYKMFHTTALIKPESGLDINKSVT---G 481

Db 237 WISVHEPLVMSFGPVVLYVMN-----GTMFLLAARTSCSTGCRKAKTSALTLS 288

Db 482 AIALCLLGLTMAFGMYINESTVIMAYLFTFINSLOGMFIETFCVLOKKVREY-GKC 540

Db 289 SFLLILLVSNWLPGLLAVVHSLIARHYLHAGCGLDGLAVLLFCVLLNADARAMPAC 348

QY 541 LRTHCCS-----GKSTESSIGSGTSGTRTPGRYS 570

Db 349 LGRKAPEARAPAGLPGAYNNMTALFEESGLIRITLGASTVSSVSARS--GRTOPDSD 406

QY 571 TGSOSRIRRMNDTVKROSSSF-----ITDDINSSASLNR 606

Db 407 QRGSRYLRD--NVLVRGSAADHTDLSLOAHAGFTDIDVDFHRDAGADSDSDLSLEE 464

QY 607 EGLLN-----NARDTSVMDLP--LNGNHGNSYSIASGE-YLSN 642

Db 465 ERSLSITSSSEEDNGRTGRFQRLPCRAGOSERLTHPKVDGNDLLSTYPALGECAAP 524

QY 643 C-----VOIDRGYNHN-----ETAL-----EKKILK-ELTSNYIPS 673

Db 525 CALQTWGSEERRLDGTSKDAANNQPPALTSGETSLGRAOROKGILKNRLQYPLVQ 584

QY 674 YLNNHRSSEONRLMKLVNNGSGR-----EDDAIYLDATSPNHESGLLEIH----- 725

Db 585 TRGAPELSMCRAATLIGRAVPAASGRIVAGGTGSLQSPASRSRREGQDLRLQLSR 644

QY 726 ---EESDAPLPP--RYVSTE-----NHOPHHYTRRRIPODSESFFPLLTNEH 769

Db 645 ERLEAPAPVLRPLSRGSOECMDAARGLEPKDRGSTLPRQPPRPPGAMAGREGSRD 704

QY 770 TEDLOSPPRDSLYTSMETLAVATSESVTSTQTEPPACGDAEDYVYKSMPLGSRNH 829

Db 705 ALDIGAP-REWL-----STLPPRRTRDLD----- 728

QY 830 VHQHTYYQLGROSSG 846

Db 729 -HSPHLCPLPSGNSG 744

RESULT 14

US-08-110-286A-6

Sequence 6, Application US/08110286A
Patent No. 5728545

GENERAL INFORMATION:

APPLICANT: Perrin, Marilyn H.

APPLICANT: Chen, Ruoping

APPLICANT: Lewis, Kathy A.

APPLICANT: Vale Jr., Wylie W.

TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSER: Pretty, Schroeder, Brueggemann & Clark

STREET: 444 South Flower Street, Suite 2000

CITY: Los Angeles

STATE: CA

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/110, 286A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/079,320

FILING DATE: 18-JUN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Reiter, Stephen E.

REGISTRATION NUMBER: 31,192

REFERENCE/DOCKET NUMBER: P41 9439

TELEPHONE: 619-546-4737

TELEFAX: 619-546-9392

INFORMATION FOR SEQ ID NO: 6:


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-286A-6

```

```

Query Match          6.4%; Score 290.5; DB 1; Length 415;
Best Local Similarity 28.6%; Pred. No. 9,8e-18;
Matches 90; Conservative 54; Mismatches 136; Indels 35; Gaps 12;

```

```

QY 248 GCRLLFTNKTHTTQSCN---HLTNFAVLMAHVEVKHSDAVHDLDDVITWVIGILSLVC 303
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 74 GVRNTNTNGYRCLANGSMAARVNSSECELINEEKKSKVHVHVAIIVLYLGHCSLVA 133
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 304 LLICITFCFEGRLQSDRNTIHNKLCISLFAELLFLIGINRDP-----IA-CAVFA 356
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 134 LVAFAVFLRLRSIRCLRNIHNNL-ISAFLRNATFVVQLTVSEVHQSNAWCRLYV 192
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 357 ALHFFFLAFTWMPLEGVQVLYIMLVEVFESESRKRYFLVGYGMPALIVANSAVDYR 416
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 193 AANYNFHVNFEMFGEGCLHTAIVLYSTDRLKRMFVCIGWVFPPII-VAMAIGKL 251
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 417 SYGTDKVCWL-RLDYFETMSFTGPATLIMLVIFLGIALYKMFHTALIKPE---SGC 471
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 252 HYDNEK-CMFGKRPVYTDYIYGPMILVLLINFIPL-----FNIVRLMTKLAST 303
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 472 LDNI--KSWVIGAILCLGLTWAFGLMYNE----STVIMAYLFTTNSLQGMFTI 524
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 304 SETIOYRKAVKATLVLLPLGLITYM--LFEVNGEDEVSRYVETIYNSFLESFQGFVS 361
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 525 FHCYLOKRYKKEYGK 539
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 362 FYCFLNSEVRSAIRK 376
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 15
US-08-981-189B-10
; Sequence 10, Application US/08981189B
; Patent No. 6214797
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UROCORTIN PEPTIDES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITCH, EVEN, TABIN & FLANNERY
; STREET: 120 S. LaSalle Street, Suite 1600
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,189B
; FILING DATE: 10-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/028,144
; FILING DATE: 13-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,223
; FILING DATE: 11-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Schumann, James J.
; REGISTRATION NUMBER: 20,856
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858-552-1311
; TELEFAX: 858-552-0095
; INFORMATION FOR SEQ ID NO: 10:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..415
; OTHER INFORMATION: /note="Human Pituitary
; OTHER INFORMATION: CRF-Receptor-R1"
; PUBLICATION INFORMATION:
; AUTHORS: Perrin, Marilyn H
; AUTHORS: Donaldson, Cynthia J
; AUTHORS: Chen, Ruoping
; AUTHORS: Lewis, Kathy A
; AUTHORS: Vale, Wylie W
; TITLE: Cloning and Functional Expression of a Rat
; TITLE: Brain Corticotropin Releasing Factor (CRF)
; TITLE: Receptor
; JOURNAL: Endocrinology
; VOLUME: 133
; ISSUE: 6
; PAGES: 3058-3061
; DATE: 1993
US-08-981-189B-10

```

```

Query Match          6.4%; Score 290.5; DB 4; Length 415;
Best Local Similarity 28.6%; Pred. No. 9,8e-18;
Matches 90; Conservative 54; Mismatches 136; Indels 35; Gaps 12;

```

```

QY 248 GCRLLFTNKTHTTQSCN---HLTNFAVLMAHVEVKHSDAVHDLDDVITWVIGILSLVC 303
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 74 GVRNTNTNGYRCLANGSMAARVNSSECELINEEKKSKVHVHVAIIVLYLGHCSLVA 133
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 304 LLICITFCFEGRLQSDRNTIHNKLCISLFAELLFLIGINRDP-----IA-CAVFA 356
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 134 LVAFAVFLRLRSIRCLRNIHNNL-ISAFLRNATFVVQLTVSEVHQSNAWCRLYV 192
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 357 ALHFFFLAFTWMPLEGVQVLYIMLVEVFESESRKRYFLVGYGMPALIVANSAVDYR 416
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 193 AANYNFHVNFEMFGEGCLHTAIVLYSTDRLKRMFVCIGWVFPPII-VAMAIGKL 251
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 417 SYGTDKVCWL-RLDYFETMSFTGPATLIMLVIFLGIALYKMFHTALIKPE---SGC 471
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 252 HYDNEK-CMFGKRPVYTDYIYGPMILVLLINFIPL-----FNIVRLMTKLAST 303
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 472 LDNI--KSWVIGAILCLGLTWAFGLMYNE----STVIMAYLFTTNSLQGMFTI 524
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 304 SETIOYRKAVKATLVLLPLGLITYM--LFEVNGEDEVSRYVETIYNSFLESFQGFVS 361
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 525 FHCYLOKRYKKEYGK 539
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 362 FYCFLNSEVRSAIRK 376
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Search completed: December 10, 2002, 11:13:23
Job time : 25 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 11:09:29 ; Search time 27 Seconds
(without alignments)
3104.786 Million cell updates/sec

Title: US-09-744-226a-1

Perfect score: 4563

Sequence: 1 AEQTRNHLNMGDIYSVRAM.....KDGPPSGSKGPAHLVTSLS 872

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4456	97.7	1503	2	T18389 latrophilin-3, spl
2	4456	97.7	1571	2	T18395 latrophilin-3, spl
3	4441.5	97.3	1512	2	T18392 latrophilin-3, spl
4	4441.5	97.3	1580	2	T18407 latrophilin-3, spl
5	4401.5	96.5	1459	2	T17186 CL3AA protein - ra
6	4401.5	96.5	1527	2	T17198 CL3BA protein - ra
7	4370	95.8	1550	2	T14327 alpha-latotoxin r
8	3123.5	68.5	1273	2	T17188 CL3AC protein - ra
9	3123.5	68.5	1341	2	T17200 CL3BC protein - ra
10	3123	68.4	1230	2	T17187 CL3AB protein - ra
11	3123	68.4	1298	2	T17199 CL3BB protein - ra
12	3120.5	68.4	1274	2	T18391 latrophilin-3, spl
13	3120.5	68.4	1342	2	T18405 latrophilin-3, spl
14	3120	68.4	1231	2	T18390 latrophilin-3, spl
15	3120	68.4	1299	2	T18398 latrophilin-3, spl
16	3106	68.1	1283	2	T18394 latrophilin-3, spl
17	3106	68.1	1351	2	T18409 latrophilin-3, spl
18	3105.5	68.1	1240	2	T18393 latrophilin-3, spl
19	3105.5	68.1	1308	2	T18408 latrophilin-3, spl
20	2646.5	58.0	1341	2	T18301 latrophilin-2 (spl)
21	2646.5	58.0	1407	2	T18381 latrophilin-2 (spl)
22	2630	57.6	1354	2	T18375 latrophilin-2 (spl)
23	2630	57.6	1420	2	T18385 latrophilin-2 (spl)
24	2624	57.5	1356	2	T18367 latrophilin-2, spl
25	2624	57.5	1422	2	T18383 latrophilin-2, spl
26	2619	57.4	1384	2	T18366 latrophilin-2 (spl)
27	2619	57.4	1450	2	T18382 latrophilin-2 (spl)
28	2607.5	57.1	1369	2	T18379 latrophilin-2 (spl)
29	2607.5	57.1	1435	2	T18387 latrophilin-2 (spl)

30	2602.5	57.0	1397	2	T18377 latrophilin-2 (spl)
31	2602.5	57.0	1463	2	T18386 latrophilin-2 (spl)
32	2596.5	56.9	1399	2	T18370 latrophilin-2 (spl)
33	2596.5	56.9	1465	2	T18384 latrophilin-2 (spl)
34	2580	56.5	1412	2	T18380 latrophilin-2 (spl)
35	2580	56.5	1478	2	T18388 latrophilin-2 (spl)
36	2577	56.5	1420	2	T17158 CL2AB protein - ra
37	2554.5	56.0	1435	2	T46611 CL2AB protein - ra
38	2549.5	55.9	1463	2	T17159 CL2AC protein - ra
39	2527	55.4	1478	2	T17185 CL2BC protein - ra
40	2521	55.2	1452	2	T17157 CL2A protein - ra
41	2518.5	55.2	1487	2	T14324 alpha-latotoxin r
42	2498.5	54.8	1467	2	T17160 CL2BA protein - ra
43	2158	47.3	1466	2	T17139 CL1AA protein - ra
44	2158	47.3	1471	2	T17149 CL1BA protein - ra
45	2148.5	47.1	1467	2	T18411 latrophilin-1, bra

ALIGNMENTS

RESULT 1

T18389 latrophilin-3, splice variant abaf, brain-specific - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_Change 18-Feb-2000

C:Accession: T18389

R:Matsumita, H.; Lellanova, V.G.; Uskaryov, Y.A.

FEBS Lett. 443, 348-352, 1999

A>Title: The latrophilin family: multiply spliced G protein-coupled receptors with di

A:Reference number: T18869; MUID:99148828; PMID:10025961

A:Accession: T18389

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1503 <MAT>

A:Cross-references: EMBL:AF111085; NID:94164052; PID:94164053; PIDN:AD05321.1

C:Superfamily: alpha-latotoxin receptor, calcium-independent

C:Keywords: alternative splicing; G protein-coupled receptor

Query Match	97.7%	Score 4456;	DB 2;	Length 1503;
Best Local Similarity	92.8%	Pred. No. 1.1e-309;		
Matches 861; Conservative	3;	Mismatches 8;	Indels 56;	Gaps 2;
QY	1	AEQTRNHLNMGDIYSVAMQVLGLDYLQNLPRGKDSARSLNK-----	48	
DB	576	AEQTRNHLNMGDIYSVRAMQVLGLDYLQNLPRGKDSARSLNKQKRESCRAV	635	
QY	49	-AMVEYNNLQPOLNLMARDLTSDDLRAATMLHTFEESAFVLADMLKTDIYREMTD	107	
DB	636	QAMVETVNNLQPOLNLMARDLTSDDLRAATMLHTFEESAFVLADMLKTDIYREMTD	695	
QY	108	NKLEVARLSTEGNLEDEKPEPNMGHSTIQLSANTLQKNGRGEIRVAFVLYNNLGYL	167	
DB	696	NKLEVARLSTEGNLEDEKPEPNMGHSTIQLSANTLQKNGRGEIRVAFVLYNNLGYL	755	
QY	168	STENNSMKLGTFTALSTNSVIVNSPVYTAATIKESNSKVIYLLADPVVFYKHKQSEENFN	227	
DB	756	STENNSMKLGTFTALSTNSVIVNSPVYTAATIKESNSKVIYLLADPVVFYKHKQSEENFN	815	
QY	228	PCGSFWSKRTMYGWSYOGCRLLTNKTHHTGSCNHLTNPAVLAHVEVKSADVDL	287	
DB	816	PCGSFWSKRTMYGWSYOGCRLLTNKTHHTGSCNHLTNPAVLAHVEVKSADVDL	875	
QY	288	LDDVITWVGILLSLVCLCIFTFCFPRGLQSDRMTIHKNLCISLFAVELFLIGINRTD	347	
DB	876	LDDVITWVGILLSLVCLCIFTFCFPRGLQSDRMTIHKNLCISLFAVELFLIGINRTD	935	
QY	348	QPIACAVRAALLHFFFLAFTWFLFEGVOLYIMLVYESESRKKRYLVGSGPALIV	407	
DB	936	QPIACAVRAALLHFFFLAFTWFLFEGVOLYIMLVYESESRKKRYLVGSGPALIV	995	
QY	408	AVSAVDKRSYGTGVKRLDLYFIWSPFGPATLIMLVFLIALYKMFHFAILKP	467	

Db 996 AVSAADVRSYGTGDKVCMRLDTYFTWISFGPATLLIMLVIFLGIALYKMFHHTAILKP 1055
Qy 468 ESGCLDNKISWVIGATALLCLGLTWAFGMLYNESSTVIMAYLFTFNSIQGFIFHC 527
Db 1056 ESGCLDNKISWVIGATALLCLGLTWAFGMLYNESSTVIMAYLFTFNSIQGFIFHC 1115
Qy 528 VLOKKVREKYGKCLRTHCCSGKSTESSIGSGKTSGRTPGYSQSRIRRMNDTVRK 587
Db 1116 VLOKKVREKYGKCLRTHCCSGKSTESSIGSGKTSGRTPGYSQSRIRRMNDTVRK 1175
Qy 588 QSESSFTTGDINSSASLNR----- 606
Db 1176 QSESSFTTGDINSSASLNR----- 606
Qy 607 --EGLLNARDTSVMDTLPNGNHGNSYSIASGEYLSNCVOITDRGYNHETALEKKILK 664
Db 1236 AOGGLLNARDTSVMDTLPNGNHGNSYSIASGEYLSNCVOITDRGYNHETALEKKILK 1295
Qy 665 ELTSNVIPTSLNNHERSSSEONRLMKLVNMLGSGREDDAIVLDATSFNHEESLGLLEI 724
Db 1296 ELTSNVIPTSLNNHERSSSEONRLMKLVNMLGSGREDDAIVLDATSFNHEESLGLLEI 1355
Qy 725 HEESDAPLLPRPVYSTENHQPHTRRRIPODSESFFLLTNEHEDLOSPPHRSLSYTS 784
Db 1356 HEESDAPLLPRPVYSTENHQPHTRRRIPODSESFFLLTNEHEDLOSPPHRSLSYTS 1415
Qy 785 MPFLAAGVAATESVTSTQTPRPAKCGDAEDVYKSMPLNGSNHYNHOLHTTYQLGRSS 844
Db 1416 MPFLAAGVAATESVTSTQTPRPAKCGDAEDVYKSMPLNGSNHYNHOLHTTYQLGRSS 1475
Qy 845 DGFIVPPNKDGTPEEGSSKGPALHVTSL 872
Db 1476 DGFIVPPNKDGTPEEGSSKGPALHVTSL 1503

RESULT 2

T18395
latrophilin-3, splice variant dbaf, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18395
R:Matsumita, H.; Lellanova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A>Title: The latrophilin family: multiply spliced G protein-coupled receptors with diffe
A:Reference number: Z18869; MUID:99148828; PMID:10025961
A:Accession: T18395
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1571 <MAT>
A:Cross-references: EMBL:AF111091; NID:94164064; PID:94164065; PIDN:AAD05327.1
C:Superfamily: alpha-latotoxin receptor, calcium-independent
C:Keywords: alternative splicing; G protein-coupled receptor

Query Match 97.7%; Score 4456; DB 2; Length 1571;
Best Local Similarity 92.8%; Pred. No. 1.2e-309;
Matches 861; Conservative 3; Mismatches 8; Indels 56; Gaps 2;

Qy 1 AEOTRHLNAGDITTVSRAMDOLVGLLDVQLRNLTPGGKDSARSLNK----- 48
Db 644 AEOTRHLNAGDITTVSRAMDOLVGLLDVQLRNLTPGGKDSARSLNKLOKRRSCRAY 703
Qy 49 -AMVETVNNLLOPQALNWRDLTTSOLRAATMLHTVERSAFVLADNLKTDIVEND 107
Db 704 QAMVETVNNLLOPQALNWRDLTTSOLRAATMLHTVERSAFVLADNLKTDIVEND 763
Qy 108 NIKLEVARLSTEGNEDLKPERMGHSTQLSANTLKONGRGEIRAVAVLYNNIGPYL 167
Db 764 NIOLEVARLSTEGNEDLKPERMGHSTQLSANTLKONGRGEIRAVAVLYNNIGPYL 823
Qy 168 STENASMKLSTEGNEDLKPERMGHSTQLSANTLKONGRGEIRAVAVLYNNIGPYL 227
Db 824 STENASMKLSTEGNEDLKPERMGHSTQLSANTLKONGRGEIRAVAVLYNNIGPYL 883

Qy 228 PNCSEFMSYKRTMTGYWSTOGCRLLTNTKHTTSCNHLTNFAVLMAHEVKKSHDAVDL 287
Db 884 PNCSEFMSYKRTMTGYWSTOGCRLLTNTKHTTSCNHLTNFAVLMAHEVKKSHDAVDL 943
Qy 288 LLDVITWVGILLSLVCLLCIFTECFRGLQSORNTIHKNLCLSLVAELFLIGINRND 347
Db 944 LLDVITWVGILLSLVCLLCIFTECFRGLQSORNTIHKNLCLSLVAELFLIGINRND 1003
Qy 348 QPACAVFALLHFFLAFTWMELEGVOLYIMLVFESEHSRRRYFLVYGPMALY 407
Db 1004 QPACAVFALLHFFLAFTWMELEGVOLYIMLVFESEHSRRRYFLVYGPMALY 1063
Qy 408 AVSAADVRSYGTGDKVCMRLDTYFTWISFGPATLLIMLVIFLGIALYKMFHHTAILKP 467
Db 1064 AVSAADVRSYGTGDKVCMRLDTYFTWISFGPATLLIMLVIFLGIALYKMFHHTAILKP 1123
Qy 468 ESGCLDNKISWVIGATALLCLGLTWAFGMLYNESSTVIMAYLFTFNSIQGFIFHC 527
Db 1124 ESGCLDNKISWVIGATALLCLGLTWAFGMLYNESSTVIMAYLFTFNSIQGFIFHC 1183
Qy 528 VLOKKVREKYGKCLRTHCCSGKSTESSIGSGKTSGRTPGYSQSRIRRMNDTVRK 587
Db 1184 VLOKKVREKYGKCLRTHCCSGKSTESSIGSGKTSGRTPGYSQSRIRRMNDTVRK 1243
Qy 588 QSESSFTTGDINSSASLNR----- 606
Db 1244 QSESSFTTGDINSSASLNR----- 606
Qy 607 --EGLLNARDTSVMDTLPNGNHGNSYSIASGEYLSNCVOITDRGYNHETALEKKILK 664
Db 1304 AOGGLLNARDTSVMDTLPNGNHGNSYSIASGEYLSNCVOITDRGYNHETALEKKILK 1363
Qy 665 ELTSNVIPTSLNNHERSSSEONRLMKLVNMLGSGREDDAIVLDATSFNHEESLGLLEI 724
Db 1364 ELTSNVIPTSLNNHERSSSEONRLMKLVNMLGSGREDDAIVLDATSFNHEESLGLLEI 1423
Qy 725 HEESDAPLLPRPVYSTENHQPHTRRRIPODSESFFLLTNEHEDLOSPPHRSLSYTS 784
Db 1424 HEESDAPLLPRPVYSTENHQPHTRRRIPODSESFFLLTNEHEDLOSPPHRSLSYTS 1483
Qy 785 MPFLAAGVAATESVTSTQTPRPAKCGDAEDVYKSMPLNGSNHYNHOLHTTYQLGRSS 844
Db 1484 MPFLAAGVAATESVTSTQTPRPAKCGDAEDVYKSMPLNGSNHYNHOLHTTYQLGRSS 1543
Qy 845 DGFIVPPNKDGTPEEGSSKGPALHVTSL 872
Db 1544 DGFIVPPNKDGTPEEGSSKGPALHVTSL 1571

RESULT 3

T18392
latrophilin-3, splice variant abdf, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18392
R:Matsumita, H.; Lellanova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A>Title: The latrophilin family: multiply spliced G protein-coupled receptors with di
A:Reference number: Z18869; MUID:99148828; PMID:10025961
A:Accession: T18392
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1512 <MAT>
A:Cross-references: EMBL:AF111088; NID:94164058; PID:94164059; PIDN:AAD05324.1
C:Superfamily: alpha-latotoxin receptor, calcium-independent
C:Keywords: alternative splicing; G protein-coupled receptor

Query Match 97.3%; Score 4441.5; DB 2; Length 1512;
Best Local Similarity 91.9%; Pred. No. 1.2e-308;
Matches 861; Conservative 3; Mismatches 8; Indels 65; Gaps 3;

Qy 1 AEOTRHLNAGDITTVSRAMDOLVGLLDVQLRNLTPGGKDSARSLNK----- 48
Db 1544 AEOTRHLNAGDITTVSRAMDOLVGLLDVQLRNLTPGGKDSARSLNK----- 48

```

Db 576 AEQTRHNLNAGDITTSYVRANDOLVGLLDVQRLMLTPGKXDSAAARSINKLOKRERSGRAY 635
QY 49 -AMETVNNLLOPQALNAMEDLTSSDQRAATMLLHTVEESAVFLADNLKTDIVREND 107
Db 636 QAAVEVNNLLOPQALNAMEDLTSSDQRAATMLLHTVEESAVFLADNLKTDIVREND 695
QY 108 NIKLEVARLSTEGNLEDLKEPENMGHSTQLSANTLKONGRGEIRVAFLVNLGPLY 167
Db 696 NIKLEVARLSTEGNLEDLKEPENMGHSTQLSANTLKONGRGEIRVAFLVNLGPLY 755
QY 168 STENASMKLGTEALSTNHSYIVNSPYITAAINKEFSNKYVLADPVVFTYKHIOSEENFN 227
Db 756 STENASMKLGTEALSTNHSYIVNSPYITAAINKEFSNKYVLADPVVFTYKHIOSEENFN 815
QY 228 PNCSEFMSYSKRTMTGWSGCRLLTNTKHTTSCSNHLTNFAVLMAHVEVKHSDAVHDL 287
Db 816 PNCSEFMSYSKRTMTGWSGCRLLTNTKHTTSCSNHLTNFAVLMAHVEVKHSDAVHDL 875
QY 288 LLDVITWVGILSLVCLLICTFFCFEFGIQSDRNTIHKMLCISLVEAELLFLGINKRD 347
Db 876 LLDVITWVGILSLVCLLICTFFCFEFGIQSDRNTIHKMLCISLVEAELLFLGINKRD 935
QY 348 QPIACAVFAALLHFFFLAFTWMLFEGVOLYIMLVEVESEHSRRKRYLVGYGMPALIV 407
Db 936 QPIACAVFAALLHFFFLAFTWMLFEGVOLYIMLVEVESEHSRRKRYLVGYGMPALIV 995
QY 408 AVSAADYRSYGTDKYCMRLDPTWTFSGPATLLIIMLVIFLGIALYKMFHTALIKR 467
Db 996 AVSAADYRSYGTDKYCMRLDPTWTFSGPATLLIIMLVIFLGIALYKMFHTALIKR 1055
QY 468 ESCGLDN-----IKSWAIGAILLCLLGLTMAFGIMYINESTVIMAYLFTIFNSIQ 518
Db 1056 ESCGLDNINEDNRPFISKVIGAILLCLLGLTMAFGIMYINESTVIMAYLFTIFNSIQ 1115
QY 519 GMFTIFHCYLOKKVREYKCLRTGCCSGKSTRESSIGSKTSGSRTPGRYSTGQSRRIR 578
Db 1116 GMFTIFHCYLOKKVREYKCLRTGCCSGKSTRESSIGSKTSGSRTPGRYSTGQSRRIR 1175
QY 579 RMNDVVRKOSSESTTGDINSASLNR----- 606
Db 1176 RMNDVVRKOSSESTTGDINSASLNRGAMANHLISNALLRPHGTNNPYNTLLGEPAYC 1235
QY 607 -----EGLNNARDTSYMDTLPLNGNNGNSYSIASGEYLSNCVOIIDRGYNHNE 655
Db 1236 NNSVSVMYNAQEBGLNNARPTSYMDTLPLNGNNGNSYSIASGEYLSNCVOIIDRGYNHNE 1295
QY 656 TALEKKILKELTSNYIPSYLNNHERSESONRNLMNKLNNVLSGSGEDDAIVLDDATSPFNH 715
Db 1296 TALEKKILKELTSNYIPSYLNNHERSESONRNLMNKLNNVLSGSGEDDAIVLDDATSPFNH 1355
QY 716 EESLGLLELTHEESDAPLLPRVYSTENHOPNHTTTRRIIPODHSESPFLLTNEHTEDLOS 775
Db 1356 EESLGLLELTHEESDAPLLPRVYSTENHOPNHTTTRRIIPODHSESPFLLTNEHTEDLOS 1415
QY 776 PHNDSLYTSMPTLAGVAALIESVTTSTQTEPRPAKCGDAEDVYKXSPNLSGRNHVQLHT 835
Db 1416 PHNDSLYTSMPTLAGVAALIESVTTSTQTEPRPAKCGDAEDVYKXSPNLSGRNHVQLHT 1475
QY 836 YVOLGRSSDGFIVPNNKDGTPREGSSKGPANLVTSLS 872
Db 1476 YVOLGRSSDGFIVPNNKDGTPREGSSKGPANLVTSLS 1512

RESULT 4
T18407
latrophilin-3, splice variant bdbf, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18407
R:Matsumita, H.; Ielanova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with differ
A:Reference number: Z18869; MUID:99148828; PMID:10025961

```

```

A:Accession: T18407
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1580 <Mat>
A:Cross-references: EMBL:AF111094; NID:94164070; PID:94164071; PIDN:AAD05330.1
C:Superfamily: alpha-latotoxin receptor, calcium-independent
C:Keywords: alternative splicing; G protein-coupled receptor

Query Match          97.3%; Score 4441.5; DB 2; Length 1580;
Best Local Similarity 91.9%; Pred. No. 1.3e-308;
Matches 861; Conservative 3; Mismatches 8; Indels 65; Gaps 3;

QY 1 AEQTRHNLNAGDITTSYVRANDOLVGLLDVQRLMLTPGKXDSAAARSINKLOKRERSGRAY 48
Db 644 AEQTRHNLNAGDITTSYVRANDOLVGLLDVQRLMLTPGKXDSAAARSINKLOKRERSGRAY 703
QY 49 -AMETVNNLLOPQALNAMEDLTSSDQRAATMLLHTVEESAVFLADNLKTDIVREND 107
Db 704 QAAVEVNNLLOPQALNAMEDLTSSDQRAATMLLHTVEESAVFLADNLKTDIVREND 763
QY 108 NIKLEVARLSTEGNLEDLKEPENMGHSTQLSANTLKONGRGEIRVAFLVNLGPLY 167
Db 764 NIKLEVARLSTEGNLEDLKEPENMGHSTQLSANTLKONGRGEIRVAFLVNLGPLY 823
QY 168 STENASMKLGTEALSTNHSYIVNSPYITAAINKEFSNKYVLADPVVFTYKHIOSEENFN 227
Db 824 STENASMKLGTEALSTNHSYIVNSPYITAAINKEFSNKYVLADPVVFTYKHIOSEENFN 883
QY 228 PNCSEFMSYSKRTMTGWSGCRLLTNTKHTTSCSNHLTNFAVLMAHVEVKHSDAVHDL 287
Db 884 PNCSEFMSYSKRTMTGWSGCRLLTNTKHTTSCSNHLTNFAVLMAHVEVKHSDAVHDL 943
QY 288 LLDVITWVGILSLVCLLICTFFCFEFGIQSDRNTIHKMLCISLVEAELLFLGINKRD 347
Db 944 LLDVITWVGILSLVCLLICTFFCFEFGIQSDRNTIHKMLCISLVEAELLFLGINKRD 1003
QY 348 QPIACAVFAALLHFFFLAFTWMLFEGVOLYIMLVEVESEHSRRKRYLVGYGMPALIV 407
Db 1004 QPIACAVFAALLHFFFLAFTWMLFEGVOLYIMLVEVESEHSRRKRYLVGYGMPALIV 1063
QY 408 AVSAADYRSYGTDKYCMRLDPTWTFSGPATLLIIMLVIFLGIALYKMFHTALIKR 467
Db 1064 AVSAADYRSYGTDKYCMRLDPTWTFSGPATLLIIMLVIFLGIALYKMFHTALIKR 1123
QY 468 ESCGLDN-----IKSWAIGAILLCLLGLTMAFGIMYINESTVIMAYLFTIFNSIQ 518
Db 1124 ESCGLDNINEDNRPFISKVIGAILLCLLGLTMAFGIMYINESTVIMAYLFTIFNSIQ 1183
QY 519 GMFTIFHCYLOKKVREYKCLRTGCCSGKSTRESSIGSKTSGSRTPGRYSTGQSRRIR 578
Db 1184 GMFTIFHCYLOKKVREYKCLRTGCCSGKSTRESSIGSKTSGSRTPGRYSTGQSRRIR 1243
QY 579 RMNDVVRKOSSESTTGDINSASLNR----- 606
Db 1244 RMNDVVRKOSSESTTGDINSASLNRGAMANHLISNALLRPHGTNNPYNTLLGEPAYC 1303
QY 607 -----EGLNNARDTSYMDTLPLNGNNGNSYSIASGEYLSNCVOIIDRGYNHNE 655
Db 1304 NNSVSVMYNAQEBGLNNARPTSYMDTLPLNGNNGNSYSIASGEYLSNCVOIIDRGYNHNE 1363
QY 656 TALEKKILKELTSNYIPSYLNNHERSESONRNLMNKLNNVLSGSGEDDAIVLDDATSPFNH 715
Db 1364 TALEKKILKELTSNYIPSYLNNHERSESONRNLMNKLNNVLSGSGEDDAIVLDDATSPFNH 1423
QY 716 EESLGLLELTHEESDAPLLPRVYSTENHOPNHTTTRRIIPODHSESPFLLTNEHTEDLOS 775
Db 1424 EESLGLLELTHEESDAPLLPRVYSTENHOPNHTTTRRIIPODHSESPFLLTNEHTEDLOS 1483
QY 776 PHNDSLYTSMPTLAGVAALIESVTTSTQTEPRPAKCGDAEDVYKXSPNLSGRNHVQLHT 835
Db 1484 PHNDSLYTSMPTLAGVAALIESVTTSTQTEPRPAKCGDAEDVYKXSPNLSGRNHVQLHT 1543
QY 836 YVOLGRSSDGFIVPNNKDGTPREGSSKGPANLVTSLS 872

```

Db 1544 YVQGRGSSDGFIVPNNKDGTPPEGSSSGPAHLVTSL 1580

RESULT 5

T17186

CL3A protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T17186

R:Sugita, S.; Ichitchenko, K.; Khvotchev, M.; Sudhof, T.C.

submitted to the EMBL Data Library, July 1998

A:Description: CL family.

A:Reference number: Z18712

A:Accession: T17186

A:Molecule type: mRNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-1459 <SUG>

A:Cross-references: EMBL:AF081154; NID:g3695134; PID:MAC62660.1

C:Superfamily: alpha-latrotoxin receptor, calcium-independent

Query Match

Best Local Similarity 96.5%; Score 4401.5; DB 2; Length 1459;

Matches 845; Conservative 10; Mismatches 17; Indels 13; Gaps 1;

Query 1 AEQTRNHLNMGDIYVRAMDVLGDLVQLRNLTPGKDSAAKSLNK-----48

Db 575 AEQTRNHLNMGDIYVRAMDVLGDLVQLRNLTPGKDSAAKSLNKLOKRESCRAY 634

Db 49 -AMEYEVNLLQPOLANMRDLTSDQRAATMLHVEESAFVLADNLKTDIVRENTD 107

Db 635 QAMETVNNLLQPOLANMRDLTSDQRAATMLDIVEESAFVLADNLKTDIVRENTD 694

Db 108 NIKLEVARLSTEGNLEDKPEPNMGHSTIQLSANTLKONGRGEIRAFVLYNNLGPYL 167

Db 695 NIOLEVARLSTEGNLEDKPEPNMGHSTIQLSANTLKONGRGEIRAFVLYNNLGPYL 754

Db 168 STENASKKLGTALSTNHSIVNSPVITTAANKESNKVYLADVPVTFVKHIKQSEENFN 227

Db 755 STENASKKLGTAMSTNHSIVNSPVITTAANKESNKVYLADVPVTFVKHIKQSEENFN 814

Db 228 PNCFSWYSKRRTMGYSTGCRLLTNTKHTTSCNHLTFNFAVLMHVEKHSDAVHDL 287

Db 815 PNCFSWYSKRRTMGYSTGCRLLTNTKHTTSCNHLTFNFAVLMHVEKHSDAVHDL 874

Db 288 LLDVITWVGILLSLVCLLICTFECFFRGLOSDRNTIHKNCISLFAVELLFLGINRTD 347

Db 875 LLDVITWVGILLSLVCLLICTFECFFRGLOSDRNTIHKNCISLFAVELLFLGINRTD 934

Db 348 QPIACAVFAALLHFFLAFTMFLVGLVQYIMLVEVESESHSRKRYFYLVGYMPALIV 407

Db 935 QPIACAVFAALLHFFLAFTMFLVGLVQYIMLVEVESESHSRKRYFYLVGYMPALIV 994

Db 408 AVSAADVRSYGTGDKVCMRLDITFIMSFIPGPAFLIIMLVNFIIGIALYKMFHHTALIKP 467

Db 995 AVSAADVRSYGTGDKVCMRLDITFIMSFIPGPAFLIIMLVNFIIGIALYKMFHHTALIKP 1054

Db 468 ESGCLDINKSVAVIGAILLCLGLTMAFGMLYINVESTIVMAYLFTIFNSLOGMFIPLFHC 527

Db 1055 ESGCLDINKSVAVIGAILLCLGLTMAFGMLYINVESTIVMAYLFTIFNSLOGMFIPLFHC 1114

Db 528 VLOKVRKREYKCLRTKCCSGKSTRESSIGSGKTSGRTPGRYSTGSGSRIIRMMNDTVRK 587

Db 1115 VLOKVRKREYKCLRTKCCSGKSTRESSIGSGKTSGRTPGRYSTGSGSRIIRMMNDTVRK 1174

Db 588 QSESSFTTGDINSASLNREGLNNARDYSMDTLPPLNGHNSYSLASGEYLSNCVOII 647

Db 1175 QSESSFTTGDINSASLNREGLNNARDYSMDTLPPLNGHNSYSLASGEYLSNCVOII 1234

Db 648 DRGNHNETALEKKILKEIISNYIPSYLNNHRSSEONRNLMNVLNNGSGRDDAIVL 707

Db 1235 DRGNHNETALEKKILKEIISNYIPSYLNNHRSSEONRNLMNVLNNGSGRDDAIVL 1294

Qy 708 DDATSFNHEESLGLIELIHEESDAPLPRLPVYSTENHOPHHYTRRRRIQDHSSEFPPLTN 767

Db 1295 DDATSFNHEESLGLIELIHEESDAPLPRLPVYSTENHOPHHYTRRRRIQDHSSEFPPLTN 1354

Qy 768 EHTEDLOSPPHDSLYTSNPTIAGVAATESYTTSTQTERPPAKCDADVDYVYKSPNIGSR 827

Db 1355 EHTEDLOSPPHDSLYTSNPTIAGVAATESYTTSTQTERPPAKCDADVDYVYKSPNIGSR 1414

Qy 828 NHHVQHTYYVQGRGSSDGFIVPNNKDGTPPEGSSSGPAHLVTSL 872

Db 1415 NHHVQHTYYVQGRGSSDGFIVPNNKDGTPPEGSSSGPAHLVTSL 1459

RESULT 6

T17198

CL3A protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T17198

R:Sugita, S.; Ichitchenko, K.; Khvotchev, M.; Sudhof, T.C.

submitted to the EMBL Data Library, July 1998

A:Description: CL family.

A:Reference number: Z18712

A:Accession: T17198

A:Molecule type: mRNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-1527 <SUG>

A:Cross-references: EMBL:AF081157; NID:g3695140; PID:g3695141; PID:MAC62663.1

C:Superfamily: alpha-latrotoxin receptor, calcium-independent

Query Match

Best Local Similarity 96.5%; Score 4401.5; DB 2; Length 1527;

Matches 845; Conservative 10; Mismatches 17; Indels 13; Gaps 1;

Query 1 AEQTRNHLNMGDIYVRAMDVLGDLVQLRNLTPGKDSAAKSLNK-----48

Db 643 AEQTRNHLNMGDIYVRAMDVLGDLVQLRNLTPGKDSAAKSLNKLOKRESCRAY 702

Db 49 -AMEYEVNLLQPOLANMRDLTSDQRAATMLHVEESAFVLADNLKTDIVRENTD 107

Db 703 QAMETVNNLLQPOLANMRDLTSDQRAATMLDIVEESAFVLADNLKTDIVRENTD 762

Db 108 NIKLEVARLSTEGNLEDKPEPNMGHSTIQLSANTLKONGRGEIRAFVLYNNLGPYL 167

Db 763 NIOLEVARLSTEGNLEDKPEPNMGHSTIQLSANTLKONGRGEIRAFVLYNNLGPYL 822

Db 168 STENASKKLGTALSTNHSIVNSPVITTAANKESNKVYLADVPVTFVKHIKQSEENFN 227

Db 823 STENASKKLGTAMSTNHSIVNSPVITTAANKESNKVYLADVPVTFVKHIKQSEENFN 882

Db 228 PNCFSWYSKRRTMGYSTGCRLLTNTKHTTSCNHLTFNFAVLMHVEKHSDAVHDL 287

Db 883 PNCFSWYSKRRTMGYSTGCRLLTNTKHTTSCNHLTFNFAVLMHVEKHSDAVHDL 942

Db 288 LLDVITWVGILLSLVCLLICTFECFFRGLOSDRNTIHKNCISLFAVELLFLGINRTD 347

Db 943 LLDVITWVGILLSLVCLLICTFECFFRGLOSDRNTIHKNCISLFAVELLFLGINRTD 1002

Db 348 QPIACAVFAALLHFFLAFTMFLVGLVQYIMLVEVESESHSRKRYFYLVGYMPALIV 407

Db 1003 QPIACAVFAALLHFFLAFTMFLVGLVQYIMLVEVESESHSRKRYFYLVGYMPALIV 1062

Db 408 AVSAADVRSYGTGDKVCMRLDITFIMSFIPGPAFLIIMLVNFIIGIALYKMFHHTALIKP 467

Db 1063 AVSAADVRSYGTGDKVCMRLDITFIMSFIPGPAFLIIMLVNFIIGIALYKMFHHTALIKP 1122

Db 468 ESGCLDINKSVAVIGAILLCLGLTMAFGMLYINVESTIVMAYLFTIFNSLOGMFIPLFHC 527

Db 1123 ESGCLDINKSVAVIGAILLCLGLTMAFGMLYINVESTIVMAYLFTIFNSLOGMFIPLFHC 1182

Db 528 VLOKVRKREYKCLRTKCCSGKSTRESSIGSGKTSGRTPGRYSTGSGSRIIRMMNDTVRK 587

Db 1183 VLOKVRKREYKCLRTKCCSGKSTRESSIGSGKTSGRTPGRYSTGSGSRIIRMMNDTVRK 1242

```

QY 588 QSESFRTGDISSASLNREGILNNARDTSVMDTLPPLNGHNSYSIASEYLNCVOIT 647
      |||||||
Db 1243 QSSSFITGDINSASLNREGILNNARDTSVMDTLPPLNGHNSYSIASEYLNCVOIT 1302
QY 648 DRGYNHNETALEKKILKELTSNYIPSYLNHHERSSQONRMKLVNMLSGREDAIVL 707
      |||||||
Db 1303 DRGYNHNETALEKKILKELTSNYIPSYLNHHERSSQONRMKLVNMLSGREDAIVL 1362
QY 708 DDAATSTNHEESLGLLELHEESDAPLLPRVYSTENHOPHHYTRRRIPDHSSEFPPLTN 767
      |||||||
Db 1363 DDAATSTNHEESLGLLELHEESDAPLLPRVYSTENHOPHHYTRRRIPDHSSEFPPLTN 1422
QY 768 EHEDEDQSPRDSLYTSMPTLAVAAFTSTSTOTEPRAKGDADVYKSGMPNLGSR 827
      |||||||
Db 1423 EHEDEDQSPRDSLYTSMPTLAVAAFTSTSTOTEPRAKGDADVYKSGMPNLGSR 1482
QY 828 NHVHQLHTYYQLGRGSSDGFIVPNKDGTPPESSSGPAHLVTSL 872
      |||||||
Db 1483 NHVHQLHTYYQLGRGSSDGFIVPNKDGTPPESSSGPAHLVTSL 1527

```

RESULT 7

```

T14327
alpha-latrotoxin receptor 3, calcium-independent - rat
N:Alternate names: G-protein coupled receptor
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #ext_change 18-Feb-2000
C:Accession: T14327
R:Ichchenko, K.A.; Biltner, M.A.; Krasnoperov, V.; Little, A.R.; Chepuray, O.; Holz, R.
J. Biol. Chem. 274, 5491-5498, 1999
A:Title: A novel ubiquitously expressed alpha-latrotoxin receptor is a member of the C1H
A:Reference number: Z17983; MIMD:99150330; PMID:10026162
A:Accession: T14327
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1550 <ICH>
A:Cross-references: EMBL:AF063103; NID:93882980; PID:93882981; PIDN:AAC77816.1
C:Genetics:
A:Gene: C1RL-3
C:Superfamily: alpha-latrotoxin receptor, calcium-independent

```

```

Query Match          95.8%; Score 4370; DB 2; Length 1550;
Best Local Similarity 93.1%; Pred. No. 1,7e-303;
Matches 845; Conservative 10; Mismatches 17; Indels 36; Gaps 3;

QY 1 AEOTRNHLNAGDITTSYVRAMDQVLGLDVQLRMLTPGKGDSAAARSINK----- 48
      |||||||
Db 643 AEOTRNHLNAGDITTSYVRAMDQVLGLDVQLRMLTPGKGDSAAARSINKLQKRRCRAVY 702
QY 49 -AMVETVNNLLOQALNAMDLTTSQDLRAATMLLTVEESAVVLADNLKTDIVERND 107
      |||||||
Db 703 QAAVETVNNLLOQALNAMDLTTSQDLRAATMLLTVEESAVVLADNLKTDIVERND 762
QY 108 NIKLEVARLSTEGNLDLKFPEMNGHSTTQLSANTLKQGRNGEIRVAFVLNNIGPYL 167
      |||||||
Db 763 NIKLEVARLSTEGNLDLKFPEMNGHSTTQLSANTLKQGRNGEIRVAFVLNNIGPYL 822
QY 168 STENASMKLQTEALSTNHSIVNSPVITAAINKEFSNKVYLADPVYFTVKHIQSEENFN 227
      |||||||
Db 823 STENASMKLQTEALSTNHSIVNSPVITAAINKEFSNKVYLADPVYFTVKHIQSEENFN 882
QY 228 PNCSEFSYSKRTMTGYWSVQGCRLTNTKTHTCSCNHLTNFVLAHVEVKSADAVHL 287
      |||||||
Db 883 PNCSEFSYSKRTMTGYWSVQGCRLTNTKTHTCSCNHLTNFVLAHVEVKSADAVHL 942
QY 288 LLDVITWVGLISLVCLLCIFTCFCEFGLOSDFRNTIHKMLCISLVAELFLIGINRTD 347
      |||||||
Db 943 LLDVITWVGLISLVCLLCIFTCFCEFGLOSDFRNTIHKMLCISLVAELFLIGINRTD 1002
QY 348 QPIACAVPALALHFFFLAAFTMNLGEGVOLYLMLVEFESEHSRRRYFLVGYGMPALIV 407
      |||||||
Db 1003 QPIACAVPALALHFFFLAAFTMNLGEGVOLYLMLVEFESEHSRRRYFLVGYGMPALIV 1062

```

```

QY 408 AVSAADYRSYGTDKYKWLRLDPYFTMSFTGPATLLIMLVFLGIALYKMFHPTILKP 467
      |||||||
Db 1063 AVSAADYRSYGTDKYKWLRLDPYFTMSFTGPATLLIMLVFLGIALYKMFHPTILKP 1122
QY 468 EESCLDN-----IKSVVIGAILCLLGLTMAFGMLYINESVYMAVLTFTNSLQ 518
      |||||||
Db 1123 EESCLDNINEDNRPPFKSVVIGAILCLLGLTMAFGMLYINESVYMAVLTFTNSLQ 1182
QY 519 GMFIFIFHCYLOKRVKREYKCLRTHCCKSGKSTRESSIGSKTSGSRTPGRYSTGSGSRIR 578
      |||||||
Db 1183 GMFIFIFHCYLOKRVKREYKCLRTHCCKSGKSTRESSIGSKTSGSRTPGRYSTGSGSRIR 1242
QY 579 RMMNDVYRKQSESSFTTGDINSSASLNK-----EGLNNAADTSVMDTLP 624
      |||||||
Db 1243 RMMNDVYRKQSESSFTTGDINSSASLNKRGSLDPCVATYTLGGLNNAADTSVMDTLP 1302
QY 625 NGNHGNSYSIASEEYLSNCVOIIDRGYNHNETALEKKILKELTSNYIPSYLNHHERSSQ 684
      |||||||
Db 1303 NGNHGNSYSIASEEYLSNCVOIIDRGYNHNETALEKKILKELTSNYIPSYLNHHERSSQ 1362
QY 685 NRMLMKLVNMLSGREDAIVLDDATSFNHEESLGLLELHEESDAPLLPRVYSTENHQ 744
      |||||||
Db 1363 NRMLMKLVNMLSGREDAIVLDDATSFNHEESLGLLELHEESDAPLLPRVYSTENHQ 1422
QY 745 PHNYTRRRIPQDHSSEFPPLTNHEHEDLOSPPRDSLYTSMPTLAVAAFTSTSTQTE 804
      |||||||
Db 1423 PHNYTRRRIPQDHSSEFPPLTNHEHEDLOSPPRDSLYTSMPTLAVAAFTSTSTQTE 1482
QY 805 PPAKKGDAEDVYKSKPNLGSRRNHVQLHTYYQLGRGSSDGFIVPNKDGTPPESSSG 864
      |||||||
Db 1483 AAKKGDADVYKSKPNLGSRRNHVQLHTYYQLGRGSSDGFIVPNKDGTPPESSSG 1542
QY 865 PAHLVTSL 872
      |||||||
Db 1543 PAHLVTSL 1550

```

RESULT 8

```

T17188
CL3AC protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 04-Mar-2000
C:Accession: T17188
R:Sugita, S.; Ichchenko, K.; Khvotchev, M.; Sudhof, T.C.
Submitted to the EMBL Data Library, July 1998
A:Description: CL family.
A:Reference number: Z18712
A:Accession: T17188
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1273 <SUG>
A:Cross-references: EMBL:AF081156; NID:93695138; PID:93695139; PIDN:AAC62662.1
C:Superfamily: alpha-latrotoxin receptor, calcium-independent

```

```

Query Match          68.5%; Score 3123.5; DB 2; Length 1273;
Best Local Similarity 93.3%; Pred. No. 1.4e-214;
Matches 609; Conservative 8; Mismatches 15; Indels 21; Gaps 3;

QY 1 AEOTRNHLNAGDITTSYVRAMDQVLGLDVQLRMLTPGKGDSAAARSINK----- 48
      |||||||
Db 575 AEOTRNHLNAGDITTSYVRAMDQVLGLDVQLRMLTPGKGDSAAARSINKLQKRRCRAVY 634
QY 49 -AMVETVNNLLOQALNAMDLTTSQDLRAATMLLTVEESAVVLADNLKTDIVERND 107
      |||||||
Db 635 QAAVETVNNLLOQALNAMDLTTSQDLRAATMLLTVEESAVVLADNLKTDIVERND 694
QY 108 NIKLEVARLSTEGNLDLKFPEMNGHSTTQLSANTLKQGRNGEIRVAFVLNNIGPYL 167
      |||||||
Db 695 NIKLEVARLSTEGNLDLKFPEMNGHSTTQLSANTLKQGRNGEIRVAFVLNNIGPYL 754
QY 168 STENASMKLQTEALSTNHSIVNSPVITAAINKEFSNKVYLADPVYFTVKHIQSEENFN 227
      |||||||

```

```
Dh 755 STENASKKLGTETAMSTHSHVIVSPVITAAINKEFSNKKVYIADPVFTVHKIKOSENFN 814
Qy 228 PNCSEFMSYKRTMTGYMSTGCGRLTTNTKTHHTSCNHLTNFAVLMHAEVKHSDAVHDL 287
Dh 815 PNCSEFMSYKRTMTGYMSTGCGRLTTNTKTHHTSCNHLTNFAVLMHAEVKHSDAVHDL 874
Qy 288 LLDVITWVGILLSLVCLLICITFTFCFPRGLOSDBNTIHKNLCLISLFAELLFLGINTRD 347
Dh 875 LLDVITWVGILLSLVCLLICITFTFCFPRGLOSDBNTIHKNLCLISLFAELLFLGINTRD 934
Qy 348 OPTACAVFAALLHFEFLAFTWMLFLEGVOLYIMLVEVEFESHSRRKFFYLGVGMPALIV 407
Dh 935 OPTACAVFAALLHFEFLAFTWMLFLEGVOLYIMLVEVEFESHSRRKFFYLGVGMPALIV 994
Qy 408 AVSAADVRSYGTGDKVCMRLDPTFIMSFIPGATLIMLVNFIPLGIALYKMFHTALIKP 467
Dh 995 AVSAADVRSYGTGDKVCMRLDPTFIMSFIPGATLIMLVNFIPLGIALYKMFHTALIKP 1054
Qy 468 ESGCLDIKSWYICAILLCLLGLTMAFGIMYINESTVIMAYLFTFINSLOGMFIETPHC 527
Dh 1055 ESGCLDIKSWYICAILLCLLGLTMAFGIMYINESTVIMAYLFTFINSLOGMFIETPHC 1114
Qy 528 VLOKKVREKYGKCLRTCCSGKSTESSIGSGKTSRPGRYSTGSGSRIIRRMNDIVRK 587
Dh 1115 VLOKKVREKYGKCLRTCCSGKSTESSIGSGKTSRPGRYSTGSGSRIIRRMNDIVRK 1174
Qy 588 QSESSFITGDIINSASINREGLNNARDTSYMDLPLNGNHG--NSYSIASGE 638
Dh 1175 QSESSFITGDIINSASINRGTMANHLMSNMLLRP-----HGTNNPYNTLLGE 1221
```

RESULT 9

```
CL3BC Protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T117200
R:Sugita, S.; Ichitchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A:Description: CL family.
A:Reference number: 218712
A:Accession: T117200
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1341 <SUG>
A:Cross-references: EMBL:AF081159; NID:G3695144; PID:G3695145; PIDN:AAC62655.1
C:Superfamily: alpha-latrotoxin receptor, calcium-independent
```

Query Match

```
Best Local Similarity 93.3%; Score 3123.5; DB 2; Length 1341;
Matches 609; Conservative 8; Mismatches 15; Indels 21; Gaps 3;
```

```
Qy 1 AEOTRNHLNMGDIYSVRADQVLGLDVLRLNLTGPGKDSAAKSLK----- 48
Dh 643 AEOTRNHLNMGDIYSVRADQVLGLDVLRLNLTGPGKDSAAKSLKLOKRRSCRAV 702
Qy 49 -AMETVNNLLQPOLANMRDLTTSQDLRAATMLLHTEESAFVLADNLKTDIVRENTD 107
Dh 703 QAMVETVNNLLQPOLANMRDLTTSQDLRAATMLDTEESAFVLADNLKTDIVRENTD 762
Qy 108 NIKLEVARLSTEGNLEDLKPEPNKGHSTIOLSANLTKONGRCEIRVAFVLYNNLGPYL 167
Dh 763 NIKLEVARLSTEGNLEDLKPEPNKGHSTIOLSANLTKONGRCEIRVAFVLYNNLGPYL 822
Qy 168 STENASKKLGTETAMSTHSHVIVSPVITAAINKEFSNKKVYIADPVFTVHKIKOSENFN 227
Dh 823 STENASKKLGTETAMSTHSHVIVSPVITAAINKEFSNKKVYIADPVFTVHKIKOSENFN 882
Qy 228 PNCSEFMSYKRTMTGYMSTGCGRLTTNTKTHHTSCNHLTNFAVLMHAEVKHSDAVHDL 287
Dh 883 PNCSEFMSYKRTMTGYMSTGCGRLTTNTKTHHTSCNHLTNFAVLMHAEVKHSDAVHDL 942
Qy 288 LLDVITWVGILLSLVCLLICITFTFCFPRGLOSDBNTIHKNLCLISLFAELLFLGINTRD 347
```

```
Dh 943 LLDVITWVGILLSLVCLLICITFTFCFPRGLOSDBNTIHKNLCLISLFAELLFLGINTRD 1002
Qy 348 OPTACAVFAALLHFEFLAFTWMLFLEGVOLYIMLVEVEFESHSRRKFFYLGVGMPALIV 407
Dh 1003 OPTACAVFAALLHFEFLAFTWMLFLEGVOLYIMLVEVEFESHSRRKFFYLGVGMPALIV 1062
Qy 408 AVSAADVRSYGTGDKVCMRLDPTFIMSFIPGATLIMLVNFIPLGIALYKMFHTALIKP 467
Dh 1063 AVSAADVRSYGTGDKVCMRLDPTFIMSFIPGATLIMLVNFIPLGIALYKMFHTALIKP 1122
Qy 468 ESGCLDIKSWYICAILLCLLGLTMAFGIMYINESTVIMAYLFTFINSLOGMFIETPHC 527
Dh 1123 ESGCLDIKSWYICAILLCLLGLTMAFGIMYINESTVIMAYLFTFINSLOGMFIETPHC 1182
Qy 528 VLOKKVREKYGKCLRTCCSGKSTESSIGSGKTSRPGRYSTGSGSRIIRRMNDIVRK 587
Dh 1183 VLOKKVREKYGKCLRTCCSGKSTESSIGSGKTSRPGRYSTGSGSRIIRRMNDIVRK 1242
Qy 588 QSESSFITGDIINSASINREGLNNARDTSYMDLPLNGNHG--NSYSIASGE 638
Dh 1243 QSESSFITGDIINSASINRGTMANHLMSNMLLRP-----HGTNNPYNTLLGE 1289
```

RESULT 10

```
CL3AB Protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T117187
R:Sugita, S.; Ichitchenko, K.; Khvotchev, M.; Sudhof, T.C.
submitted to the EMBL Data Library, July 1998
A:Description: CL family.
A:Reference number: 218712
A:Accession: T117187
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1230 <SUG>
A:Cross-references: EMBL:AF081155; NID:G3695136; PID:G3695137; PIDN:AAC62661.1
C:Superfamily: alpha-latrotoxin receptor, calcium-independent
```

Query Match

```
Best Local Similarity 91.9%; Score 3123; DB 2; Length 1230;
Matches 614; Conservative 8; Mismatches 20; Indels 26; Gaps 4;
```

```
Qy 1 AEOTRNHLNMGDIYSVRADQVLGLDVLRLNLTGPGKDSAAKSLK----- 48
Dh 575 AEOTRNHLNMGDIYSVRADQVLGLDVLRLNLTGPGKDSAAKSLKLOKRRSCRAV 634
Qy 49 -AMETVNNLLQPOLANMRDLTTSQDLRAATMLLHTEESAFVLADNLKTDIVRENTD 107
Dh 635 QAMVETVNNLLQPOLANMRDLTTSQDLRAATMLDTEESAFVLADNLKTDIVRENTD 694
Qy 108 NIKLEVARLSTEGNLEDLKPEPNKGHSTIOLSANLTKONGRCEIRVAFVLYNNLGPYL 167
Dh 695 NIKLEVARLSTEGNLEDLKPEPNKGHSTIOLSANLTKONGRCEIRVAFVLYNNLGPYL 754
Qy 168 STENASKKLGTETAMSTHSHVIVSPVITAAINKEFSNKKVYIADPVFTVHKIKOSENFN 227
Dh 755 STENASKKLGTETAMSTHSHVIVSPVITAAINKEFSNKKVYIADPVFTVHKIKOSENFN 814
Qy 228 PNCSEFMSYKRTMTGYMSTGCGRLTTNTKTHHTSCNHLTNFAVLMHAEVKHSDAVHDL 287
Dh 815 PNCSEFMSYKRTMTGYMSTGCGRLTTNTKTHHTSCNHLTNFAVLMHAEVKHSDAVHDL 874
Qy 288 LLDVITWVGILLSLVCLLICITFTFCFPRGLOSDBNTIHKNLCLISLFAELLFLGINTRD 347
Dh 875 LLDVITWVGILLSLVCLLICITFTFCFPRGLOSDBNTIHKNLCLISLFAELLFLGINTRD 934
Qy 348 OPTACAVFAALLHFEFLAFTWMLFLEGVOLYIMLVEVEFESHSRRKFFYLGVGMPALIV 407
Dh 935 OPTACAVFAALLHFEFLAFTWMLFLEGVOLYIMLVEVEFESHSRRKFFYLGVGMPALIV 994
```



```

QY 408 AVSAANDYRSYGTDKVCMRLDLYFIWSEFGPATLIMLVIFGLALYKMFHHTAIKLP 467
   |||||
DB 995 AVSAANDYRSYGTDKVCMRLDLYFIWSEFGPATLIMLVIFGLALYKMFHHTAIKLP 1054
QY 468 ESGCLDNKSWVIGATALLCLGLTFAFGIAMYINVESTVIMAYLFTFNSLQGMFIFHC 527
   |||||
DB 1055 ESGCLDNKSWVIGATALLCLGLTFAFGIAMYINVESTVIMAYLFTFNSLQGMFIFHC 1114
QY 528 VLQKVKRYKGYGCLRTHCSSGKSTESSIGSGRTSGRSTGQSOSIRRMNDTVRK 587
   |||||
DB 1115 VLQKVKRYKGYGCLRTHCSSGKSTESSIGSGRTSGRSTGQSOSIRRMNDTVRK 1174
QY 588 QSESEFTGDINSSASLNEGLLNARDTSVMDTLPNGNHGNSYSIASGEYLSNCVOIT 647
   |||||
DB 1175 QSESEFTGDINSSASLNEGLLNARDTSVMDTLPNGNHGNSYSIASGEYLSNCVOIT 1221
QY 648 DRGYNHNE 655
   ||:|
DB 1222 CHGYSTTE 1229

```

RESULT 11

```

T17199
CL3BB protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T17199
R:Subsita, S.; Ichtenenko, K.; Khvotchev, M.; Sudhof, T. C.
Submitted to the EMBL Data Library, July 1998
A:Description: CL family.
A:Reference number: 218712
A:Accession: T17199
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1298 <SUG>
A:Cross-references: EMBL:AF081158; NID:g3695142; PID:g3695143; PIDN:AAC6264.1
C:Superfamily: alpha-latrotoxin receptor, calcium-independent

```

```

Query Match      68.4%; Score 3123; DB 2; Length 1298;
Best Local Similarity 91.9%; Pred. No. 1.6e-214;
Matches 614; Conservative 8; Mismatches 20; Indels 26; Gaps 4;

QY 1 AEGTRNHLNAGDITTSVRAMDQVGLLDVQLRNLTPGKDSARSLSNK----- 48
   |||||
DB 643 AEGTRNHLNAGDITTSVRAMDQVGLLDVQLRNLTPGKDSARSLSNKQKRESCRAY 702
QY 49 -AMETVNNLQPOLNANRDLTTSQDLRAATMLLTVEESAFVLADNLKTDIVREND 107
   |||||
DB 703 QAMETVNNLQPOLNANRDLTTSQDLRAATMLLTVEESAFVLADNLKTDIVREND 762
QY 108 NIKLEVARLSTEGNLEDLKFPEKMGHSTIOLSANTLKONGRGEIRVAFVLYNNIGPYL 167
   |||||
DB 763 NIKLEVARLSTEGNLEDLKFPEKMGHSTIOLSANTLKONGRGEIRVAFVLYNNIGPYL 822
QY 168 STENASMKLGTEALSTNHSVIYNSPVITAAINKEFSNKYLLADPVVETFKHIKQSEEN 227
   |||||
DB 823 STENASMKLGTEALSTNHSVIYNSPVITAAINKEFSNKYLLADPVVETFKHIKQSEEN 882
QY 228 PNCSEFSYSKRMTGYSWGCRLLTNTKTHHTSCNHLTNFAVLAHAEVKSADAVHDL 287
   |||||
DB 883 PNCSEFSYSKRMTGYSWGCRLLTNTKTHHTSCNHLTNFAVLAHAEVKSADAVHDL 942
QY 288 LLDVITWGLISLVCLLCIFTCFPRGLQSDRNTHKNLCSLVAELFLIGINRND 347
   |||||
DB 943 LLDVITWGLISLVCLLCIFTCFPRGLQSDRNTHKNLCSLVAELFLIGINRND 1002
QY 348 QPIACAVFALLHFFELAAFTWMLFEGVOLYIMLVEFESEHSRRKRYFLVGYGMPALIV 407
   |||||
DB 1003 QPIACAVFALLHFFELAAFTWMLFEGVOLYIMLVEFESEHSRRKRYFLVGYGMPALIV 1062
QY 408 AVSAANDYRSYGTDKVCMRLDLYFIWSEFGPATLIMLVIFGLALYKMFHHTAIKLP 467
   |||||
DB 1063 AVSAANDYRSYGTDKVCMRLDLYFIWSEFGPATLIMLVIFGLALYKMFHHTAIKLP 1122

```

```

QY 468 ESGCLDNKSWVIGATALLCLGLTFAFGIAMYINVESTVIMAYLFTFNSLQGMFIFHC 527
   |||||
DB 1123 ESGCLDNKSWVIGATALLCLGLTFAFGIAMYINVESTVIMAYLFTFNSLQGMFIFHC 1182
QY 528 VLQKVKRYKGYGCLRTHCSSGKSTESSIGSGRTSGRSTGQSOSIRRMNDTVRK 587
   |||||
DB 1183 VLQKVKRYKGYGCLRTHCSSGKSTESSIGSGRTSGRSTGQSOSIRRMNDTVRK 1242
QY 588 QSESEFTGDINSSASLNEGLLNARDTSVMDTLPNGNHGNSYSIASGEYLSNCVOIT 647
   |||||
DB 1243 QSESEFTGDINSSASLNEGLLNARDTSVMDTLPNGNHGNSYSIASGEYLSNCVOIT 1289
QY 648 DRGYNHNE 655
   ||:|
DB 1290 CHGYSTTE 1297

```

RESULT 12

```

T18391
latrophilin-3, splice variant abah, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18391
R:Matsumita, H.; Lellanova, V. G.; Uskaryov, Y. A.
FEBS Lett. 443, 348-352, 1999
A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with di
A:Reference number: 218869; M0ID:99148828; PMID:10023961
A:Accession: T18391
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1274 <MAT>
A:Cross-references: EMBL:AF111087; NID:g4164056; PID:g4164057; PIDN:AAD05323.1
C:Superfamily: alpha-latrotoxin receptor, calcium-independent
C:Keywords: alternative splicing; G protein-coupled receptor

```

```

Query Match      68.4%; Score 3120.5; DB 2; Length 1274;
Best Local Similarity 93.1%; Pred. No. 2.4e-214;
Matches 608; Conservative 9; Mismatches 15; Indels 21; Gaps 3;

QY 1 AEGTRNHLNAGDITTSVRAMDQVGLLDVQLRNLTPGKDSARSLSNK----- 48
   |||||
DB 576 AEGTRNHLNAGDITTSVRAMDQVGLLDVQLRNLTPGKDSARSLSNKQKRESCRAY 635
QY 49 -AMETVNNLQPOLNANRDLTTSQDLRAATMLLTVEESAFVLADNLKTDIVREND 107
   |||||
DB 636 QAMETVNNLQPOLNANRDLTTSQDLRAATMLLTVEESAFVLADNLKTDIVREND 695
QY 108 NIKLEVARLSTEGNLEDLKFPEKMGHSTIOLSANTLKONGRGEIRVAFVLYNNIGPYL 167
   |||||
DB 696 NIKLEVARLSTEGNLEDLKFPEKMGHSTIOLSANTLKONGRGEIRVAFVLYNNIGPYL 755
QY 168 STENASMKLGTEALSTNHSVIYNSPVITAAINKEFSNKYLLADPVVETFKHIKQSEEN 227
   |||||
DB 756 STENASMKLGTEALSTNHSVIYNSPVITAAINKEFSNKYLLADPVVETFKHIKQSEEN 815
QY 228 PNCSEFSYSKRMTGYSWGCRLLTNTKTHHTSCNHLTNFAVLAHAEVKSADAVHDL 287
   |||||
DB 816 PNCSEFSYSKRMTGYSWGCRLLTNTKTHHTSCNHLTNFAVLAHAEVKSADAVHDL 875
QY 288 LLDVITWGLISLVCLLCIFTCFPRGLQSDRNTHKNLCSLVAELFLIGINRND 347
   |||||
DB 876 LLDVITWGLISLVCLLCIFTCFPRGLQSDRNTHKNLCSLVAELFLIGINRND 935
QY 348 QPIACAVFALLHFFELAAFTWMLFEGVOLYIMLVEFESEHSRRKRYFLVGYGMPALIV 407
   |||||
DB 936 QPIACAVFALLHFFELAAFTWMLFEGVOLYIMLVEFESEHSRRKRYFLVGYGMPALIV 995
QY 408 AVSAANDYRSYGTDKVCMRLDLYFIWSEFGPATLIMLVIFGLALYKMFHHTAIKLP 467
   |||||
DB 996 AVSAANDYRSYGTDKVCMRLDLYFIWSEFGPATLIMLVIFGLALYKMFHHTAIKLP 1055
QY 468 ESGCLDNKSWVIGATALLCLGLTFAFGIAMYINVESTVIMAYLFTFNSLQGMFIFHC 527

```

```
|||||
Db 1056 ESCLDNKIKSWVIGALALCLLGLTMAFGIMYNESVIMAYLFTFINSLOGMFIFFHC 1115
QY 528 VLOKKVAKKEYGKCLRTKCCSGKSTRESSIGSGKTSGRTPGRYSGTSGOSRIRRMNDVVR 587
|||||
Db 1116 VLOKKVAKKEYGKCLRTKCCSGRSTRESSIGSGKTSGRTPGRYSGTSGOSRIRRMNDVVR 1175
QY 588 QSESSFTGDISNSASINREGCLNNARDTSVMDTLPPLNGNHG--NSYSIASGE 638
|||||
Db 1176 QSESSFTGDISNSASINRGAMANHLSNMLLRP-----HGTNNPYNTLLGE 1222
|||||
RESULT 13
T18405
latrophilin-3, splice variant bbah, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18405
R:Matsushita, H.; Leljanova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with diffe
A:Reference number: Z18869; MVID:99148828; PMID:10025961
A:Accession: T18405
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1342 <MAT>
A:Cross-references: EMBL:AF11093; NID:94164068; PID:94164069; PIDN:AAD05329.1
C:Superfamily: alpha-latrototoxin receptor, calcium-independent

Query Match 68.4%; Score 3120.5; DB 2; Length 1342;
Best Local Similarity 93.1%; Pred. No. 2,6e-214;
Matches 608; Conservative 9; Mismatches 15; Indels 21; Gaps 3;

QY 1 AEOTRNHLNMGDITTSVRADOLVGLLDVQLRNLTPEGKDSAAARSLRK----- 48
Db 644 AEOTRNHLNMGDITTSVRADOLVGLLDVQLRNLTPEGKDSAAARSLRKLOKRESCAYV 703
QY 49 -AMETVNNLLQPOLNAMBRLDITSDQLRAATMLLHTVESAFYLAADNLKTDIVRENTD 107
Db 704 QAMETVNNLLQPOLNAMBRLDITSDQLRAATMLDVTESAFYLAADNLKTDIVRENTD 763
QY 108 NIKLEVARLSTEGNLEDLKPEKNGHGSTIQLSANTLKONGRGEIVAAVLYNNLCPLY 167
Db 764 NIKLEVARLSTEGNLEDLKPEKNGHGSTIQLSANTLKONGRGEIVAAVLYNNLCPLY 823
QY 168 STENASMKLCTEALSTNHSIYVNSPVTTAIIKEFSKVVYLADEVVTVVTHIKOSEENFN 227
Db 824 STENASMKLCTEALSTNHSIYVNSPVTTAIIKEFSKVVYLADEVVTVVTHIKOSEENFN 883
QY 228 PNCSEFMSYKRTMGYSTGGCRLLTTNKTHTTSCNHLTNFAVLAHAEVKHSDAVHDL 287
Db 884 PNCSEFMSYKRTMGYSTGGCRLLTTNKTHTTSCNHLTNFAVLAHAEVKHSDAVHDL 943
QY 288 LLDVITVWGLLSLVCLLICITFTFCFPRGLQSDRNTIHKMLCISLFAVELLFLGINKRTD 347
Db 944 LLDVITVWGLLSLVCLLICITFTFCFPRGLQSDRNTIHKMLCISLFAVELLFLGINKRTD 1003
QY 348 OPTACAVFALLHFFLAATFMFLVEGVOLYIMLVEFESEHSRKKFYLVGMPALIV 407
Db 1004 OPTACAVFALLHFFLAATFMFLVEGVOLYIMLVEFESEHSRKKFYLVGMPALIV 1063
QY 408 AVSAADVRSYGTDKVCMRLDITFTWSFGPATLIIIMLVIFGLALYKMFHHTAILKP 467
Db 1064 AVSAADVRSYGTDKVCMRLDITFTWSFGPATLIIIMLVIFGLALYKMFHHTAILKP 1123
QY 468 ESCLDNKIKSWVIGALALCLLGLTMAFGIMYNESVIMAYLFTFINSLOGMFIFFHC 527
Db 1124 ESCLDNKIKSWVIGALALCLLGLTMAFGIMYNESVIMAYLFTFINSLOGMFIFFHC 1183
QY 528 VLOKKVAKKEYGKCLRTKCCSGKSTRESSIGSGKTSGRTPGRYSGTSGOSRIRRMNDVVR 587
Db 1184 VLOKKVAKKEYGKCLRTKCCSGKSTRESSIGSGKTSGRTPGRYSGTSGOSRIRRMNDVVR 1243
```

```
QY 588 QSESSFTGDISNSASINREGCLNNARDTSVMDTLPPLNGNHG--NSYSIASGE 638
|||||
Db 1244 QSESSFTGDISNSASINRGAMANHLSNMLLRP-----HGTNNPYNTLLGE 1290
|||||
RESULT 14
T18390
latrophilin-3, splice variant abag, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18390
R:Matsushita, H.; Leljanova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with di
A:Reference number: Z18869; MVID:99148828; PMID:10025961
A:Accession: T18390
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1231 <MAT>
A:Cross-references: EMBL:AF11086; NID:94164054; PID:94164055; PIDN:AAD05322.1
C:Superfamily: alpha-latrototoxin receptor, calcium-independent
C:Keywords: alternative splicing; G protein-coupled receptor
```

```
Query Match 68.4%; Score 3120; DB 2; Length 1231;
Best Local Similarity 91.8%; Pred. No. 2,4e-214;
Matches 613; Conservative 9; Mismatches 20; Indels 26; Gaps 4;

QY 1 AEOTRNHLNMGDITTSVRADOLVGLLDVQLRNLTPEGKDSAAARSLRK----- 48
Db 576 AEOTRNHLNMGDITTSVRADOLVGLLDVQLRNLTPEGKDSAAARSLRKLOKRESCAYV 635
QY 49 -AMETVNNLLQPOLNAMBRLDITSDQLRAATMLLHTVESAFYLAADNLKTDIVRENTD 107
Db 636 QAMETVNNLLQPOLNAMBRLDITSDQLRAATMLDVTESAFYLAADNLKTDIVRENTD 695
QY 108 NIKLEVARLSTEGNLEDLKPEKNGHGSTIQLSANTLKONGRGEIVAAVLYNNLCPLY 167
Db 696 NIKLEVARLSTEGNLEDLKPEKNGHGSTIQLSANTLKONGRGEIVAAVLYNNLCPLY 755
QY 168 STENASMKLCTEALSTNHSIYVNSPVTTAIIKEFSKVVYLADEVVTVVTHIKOSEENFN 227
Db 756 STENASMKLCTEALSTNHSIYVNSPVTTAIIKEFSKVVYLADEVVTVVTHIKOSEENFN 815
QY 228 PNCSEFMSYKRTMGYSTGGCRLLTTNKTHTTSCNHLTNFAVLAHAEVKHSDAVHDL 287
Db 816 PNCSEFMSYKRTMGYSTGGCRLLTTNKTHTTSCNHLTNFAVLAHAEVKHSDAVHDL 875
QY 288 LLDVITVWGLLSLVCLLICITFTFCFPRGLQSDRNTIHKMLCISLFAVELLFLGINKRTD 347
Db 876 LLDVITVWGLLSLVCLLICITFTFCFPRGLQSDRNTIHKMLCISLFAVELLFLGINKRTD 935
QY 348 OPTACAVFALLHFFLAATFMFLVEGVOLYIMLVEFESEHSRKKFYLVGMPALIV 407
Db 936 OPTACAVFALLHFFLAATFMFLVEGVOLYIMLVEFESEHSRKKFYLVGMPALIV 995
QY 408 AVSAADVRSYGTDKVCMRLDITFTWSFGPATLIIIMLVIFGLALYKMFHHTAILKP 467
Db 996 AVSAADVRSYGTDKVCMRLDITFTWSFGPATLIIIMLVIFGLALYKMFHHTAILKP 1055
QY 468 ESCLDNKIKSWVIGALALCLLGLTMAFGIMYNESVIMAYLFTFINSLOGMFIFFHC 527
Db 1056 ESCLDNKIKSWVIGALALCLLGLTMAFGIMYNESVIMAYLFTFINSLOGMFIFFHC 1115
QY 528 VLOKKVAKKEYGKCLRTKCCSGKSTRESSIGSGKTSGRTPGRYSGTSGOSRIRRMNDVVR 587
Db 1116 VLOKKVAKKEYGKCLRTKCCSGRSTRESSIGSGKTSGRTPGRYSGTSGOSRIRRMNDVVR 1175
QY 588 QSESSFTGDISNSASINREGCLNNARDTSVMDTLPPLNGNHGNSYASGEYSNCVOII 647
Db 1176 QSESSFTGDISNSASINREPY-----RETSMGVKLNT-----AYQIGASE---QCQGYK 1222
QY 648 DRGYNHNE 655
|||||
```

Db 1223 CHGYSTTE 1230

RESULT 15

T18398
latrophilin-3, splice variant bbaq, brain-specific - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18398
R:Matsushita, H.; Leljanova, V.G.; Ushkaryov, Y.A.
FEBS Lett. 443, 348-352, 1999
A:Title: The latrophilin family: multiply spliced G protein-coupled receptors with different reference number: 218869; PMID:99148828; PMID:10025961
A:Accession: T18398
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1299 <MAT>
A:Cross-references: EMBL:AF111092; NID:94164066; PID:94164067; PIDN:AAD05328.1
C:Superfamily: alpha-latotoxin receptor, calcium-independent
C:Keywords: alternative splicing; G protein-coupled receptor

Query Match 68.4%; Score 3120; DB 2; Length 1299;
Best Local Similarity 91.8%; Pred. No. 2.6e-214;
Matches 613; Conservative 9; Mismatches 20; Indels 26; Gaps 4;

```
QY 1 AEQTNHNLNAGDITYSVRAMDQVGLDVLQNLTPGGKDSARSLNK----- 48
Db 644 AEQTNHNLNAGDITYSVRAMDQVGLDVLQNLTPGGKDSARSLNKQKRERSGRAY 703
QY 49 -AMVETVNNLQPOLNANRDLTSDQRAATMLHTVEESAFVLADNLTKTDIVRENTD 107
Db 704 QAMVETVNNLQPOLNANRDLTSDQRAATMLDIVEESAFVLADNLTKTDIVRENTD 763
QY 108 NIKLEVARLSTEGNLEDLKFPENMGHGSTIOLSANFLKONGRNGEIRVAEVLNNLGPYL 167
Db 764 NIOLEVARLSTEGNLEDLKFPENMGHGSTIOLSANFLKONGRNGEIRVAEVLNNLGPYL 823
QY 168 STENASMKLGTEALSTNHSVITNSPVITAAINKEFSNKVYLADPVVFTYKHITKOSENFN 227
Db 824 STENASMKLGTEAMSTNHSVITNSPVITAAINKEFSNKVYLADPVVFTYKHITKOSENFN 883
QY 228 PNCSEFMSYKSRMTGMYSTOGCRLLTNTKHTTSCNHLTNFAVLAHAEVKHSDAVHDL 287
Db 884 PNCSEFMSYKSRMTGMYSTOGCRLLTNTKHTTSCNHLTNFAVLAHAEVKHSDAVHDL 943
QY 288 LLDVITWVGILLISVCLLICIFTECFEFGLOSDFNTIHKNLICISLFAELLFLIGINRTD 347
Db 944 LLDVITWVGILLISVCLLICIFTECFEFGLOSDFNTIHKNLICISLFAELLFLIGINRTD 1003
QY 348 QPIACAVFAALLHFFFLAFTMFLFEGVQLYIMLVEFESEHSRKKFYLVGCGMPALIV 407
Db 1004 QPIACAVFAALLHFFFLAFTMFLFEGVQLYIMLVEFESEHSRKKFYLVGCGMPALIV 1063
QY 408 AVSAADVRSYGTDKVCMRLDITYFMSFIGPATLIMLVFIIGALYKMFHHTAILKP 467
Db 1064 AVSAADVRSYGTDKVCMRLDITYFMSFIGPATLIMLVFIIGALYKMFHHTAILKP 1123
QY 468 ESGCLDNKISWVIGAILLCLGLTMAFGMLYINESTVIMAYLFTIFNSLQGMFIIFHC 527
Db 1124 ESGCLDNKISWVIGAILLCLGLTMAFGMLYINESTVIMAYLFTIFNSLQGMFIIFHC 1183
QY 528 VLQKKVREYKCKLRTGCCSGKSTSSIGSGKTSRTPGRYSTGQSQRIRRMNDYVRK 587
Db 1184 VLQKKVREYKCKLRTGCCSGKSTSSIGSGKTSRTPGRYSTGQSQRIRRMNDYVRK 1243
QY 588 QSESSFITGDIINSASLREGILNNAADTSVMDTLPLNGHNSYSIASGEYLSNCGYQII 647
Db 1244 QSESSFITGDIINSASLREPI-----RETSMGVKLNI-----AYQIGASF---QCQGYK 1290
QY 648 DRGYNHNE 655
Db 1291 CHGYSTTE 1298
```

Search completed: December 10, 2002, 11:13:10
Job time : 46 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 10:57:49 ; Search time 15 Seconds
(without alignments)
2411.157 Million cell updates/sec

Title: US-09-744-226A-1
Perfect score: 4563
Sequence: 1 AECFRNLNMGDIYSVRAM.....KDGTPPGSSSKGAHLVTSLS 872

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	762	16.7	835	1	CD97_HUMAN
2	749.5	16.4	931	1	EMR1_MOUSE
3	715.5	15.7	886	1	EMR1_HUMAN
4	700.5	15.4	3034	1	CLR1_MOUSE
5	692	15.2	3014	1	CLR1_HUMAN
6	637.5	14.0	2144	1	CLR2_RAT
7	636.5	13.9	2923	1	CLR2_HUMAN
8	623.5	13.7	2920	1	CLR2_MOUSE
9	578.5	12.7	1522	1	BAI3_HUMAN
10	520	11.4	3313	1	CLR3_RAT
11	518.5	11.4	3301	1	CLR3_MOUSE
12	507.5	11.1	3312	1	CLR3_HUMAN
13	505.5	11.1	1572	1	BAI2_HUMAN
14	496.5	10.9	1584	1	BAI1_HUMAN
15	483	10.6	3579	1	STAN DROME
16	302	6.6	441	1	DIHR_ACHDO
17	290.5	6.4	415	1	CRF1_MOUSE
18	290.5	6.4	415	1	CRF1_RAT
19	279.5	6.1	415	1	CRF1_SHEEP
20	273	6.0	444	1	CRF1_HUMAN
21	270.5	5.9	420	1	CRF1_CHICK
22	268	5.9	431	1	CRF2_XENLA
23	261.5	5.7	431	1	CRF2_MOUSE
24	260.5	5.7	411	1	CRF2_RAT
25	256	5.6	415	1	DIHR_XENLA
26	250.5	5.5	415	1	CRF1_XENLA
27	245.5	5.4	466	1	GIPIR_HUMAN
28	241.5	5.3	466	1	PACR_HUMAN
29	240	5.3	461	1	CGRR_HUMAN
30	238.5	5.2	496	1	PACR_HUMAN
31	238.5	5.2	411	1	CRF2_MOUSE
32	238.5	5.2	513	1	PACR_BOVIN
33	237	5.2	489	1	GLP1_MOUSE

34	236	5.2	464	1	CGRR_RAT	063118 rattus norv
35	235	5.2	478	1	CALR_CAVPO	008893 cavia porce
36	231.5	5.1	455	1	GIPIR_RAT	P43219 rattus norv
37	230	5.0	445	1	SCRC_RABIT	046502 oryctolagus
38	228	5.0	440	1	SCRC_HUMAN	P47872 homo sapien
39	225	4.9	457	1	VIPR_HUMAN	P32241 homo sapien
40	225	4.9	591	1	PTRR_MOUSE	P41593 mus musculu
41	224.5	4.9	449	1	SCRC_RAT	P23811 rattus norv
42	223.5	4.9	463	1	GLP1_RAT	P32301 rattus norv
43	223	4.9	490	1	CALR_HUMAN	P30988 homo sapien
44	223	4.9	644	1	Y044_CAEEL	Q09460 caenorhabdi
45	222.5	4.9	463	1	GLP1_HUMAN	P43220 homo sapien

ALIGNMENTS

```

RESULT 1
CD97_HUMAN          STANDARD:      PRT:      835 AA.
AC  P48960:          01-FEB-1996 (Rel. 33, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Leucocyte antigen CD97 precursor.
GN  CD97.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  MEDLINE=95363161; PubMed=7636245;
RA  Hamann J., Eichler W., Hamann D., Kerstens H.M.J., Poddighe P.J.,
RA  Hoovers J.M.N., Hartmann J.M., Straus M., van Lier R.A.W.;
RT  "Expression cloning and chromosomal mapping of the leukocyte
RT  activation antigen CD97, a new seven-span transmembrane molecule of
RT  the secretin receptor superfamily with an unusual extracellular
RT  domain.";
RL  J. Immunol. 155:1942-1950(1995).
RN  [2]
RP  SEQUENCE FROM N.A.
RP  TISSUE=ForeSkin;
RX  MEDLINE=96230339; PubMed=8786105;
RA  Hamann J., Hartmann B., van Lier R.A.W.;
RT  "Structure of the human CD97 gene: exon shuffling has generated a new
RT  type of seven-span transmembrane molecule related to the secretin
RT  receptor superfamily.";
RL  Genomics 32:144-147(1996).
RN  [3]
RP  REVISIONS.
RA  Hamann J.;
RL  Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: COULD BE A RECEPTOR POTENTIALLY INVOLVED IN BOTH
CC  ADHESION AND SIGNALING PROCESSES EARLY AFTER LEUKOCYTE ACTIVATION.
CC  -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC  -1- SIMILARITY: CONTRAINS 5 EGF-LIKE DOMAINS.
CC  -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC  -1- SIMILARITY: CONTRAINS 1 GPS DOMAIN.
CC  -1- DATABASE: NAME=PROV; NOTE=CD guide CD97 entry;
CC  WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd97.htm".
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC  or send an email to license@isb-sdb.ch).
CC  -----
CC  EMBL: X84700; CAA59173.1;
CC  EMBL: X94630; CAA64333.1;
CC  EMBL: X94631; CAA64333.1; JOINED.

```

FT	DISUFID	171	186	BY SIMILARITY.
FT	DISUFID	188	207	BY SIMILARITY.
FT	DISUFID	213	226	BY SIMILARITY.
FT	DISUFID	220	235	BY SIMILARITY.
FT	DISUFID	237	256	BY SIMILARITY.
FT	CARBOHYD	33	33	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	38	38	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	108	108	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	203	203	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	371	371	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	406	406	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	413	413	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	453	453	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	520	520	N-LINKED (GLCNAC. . .)
FT	CONFLICT	121	213	MISSING (IN REF. 1).
SO	SEQUENCE	835 AA;	91941 MW;	050672E7A4C12A9E CRC64;
Query March 16.7%; Score 762; DB 1; Length 835;				
Best local Similarity 32.9%; Pred. No. 6, 1e-42;				
Matches 189; Conservative 108; Mismatches 213; Indels 64; Gaps 18;				
QY	30	QLRNLTPEGKDSARSLSKAMVEYEVNLLQ-PQALNARDLTTSQDLRAATMLHTVEES	88	
DB	281	KVQDLGRDSKTSASAEVTIQNIYKLVDELMEAPGDVEA--LAPVRHILATQVLSNLEDI	337	
QY	89	AFVLADNLKTDIYRENTDNIKLEVARLSTGCSNLEDLKFPENMGHG-STQLSANTLKON	147	
DB	338	MRIIAKSLPKRPFTYISPSNTELT-MIOERGD-----KNVTMGOSARMKLMWAVAA	389	
QY	148	GRN--GEIRVAFVLYNNGPYLSSTENASMKGTLEASTNHSVYNSPVY-----TAAI	198	
DB	390	GAEDPPRAVAGILSTIONMTLLA--NMSLNHKKQALEETIESS--TRGVQLRLSLAV	445	
QY	199	NKEE---SNKYLLADPVVFYFKHIKQSE-----ENFNP-----CSFWSYSKRRTM	240	
DB	446	NSIFLSHNNTKELMSPLIFAFSHLESSDGEGRDPAADVMPGRQELLCAFWK--SDSDR	504	
QY	241	TGVWSTGCRLLTNTKTHHTTSCNHLTFEAVLMAHVEYKHSDAVHDLIDVITWVGILLS	300	
DB	505	GGHWATEVCQVILGKNGSTGCCSHLSFSTILMAHYDE-----DMKLTITRIGVALS	558	
QY	301	LVCLLICITCFPCFRGLGSDRNTHKNICISLFAVELLFLGINKRTDPIA--CAVFAAL	358	
DB	559	LECLLICITFLFLVLRPIGSRFTTHLHLCICLFVGSTFLGIENEGGQVGLRCLVAGL	618	
QY	359	LHFEFLAFTMPLFEGVOLYIMLVEVEFSEHSRRKYFYLVGCMPLALIVASAADVRSY	418	
DB	619	LHYCFELAFCMSSLEGELEFLVYVRVGGQSLSTRMCLITIGYVPLLIVGSAALYSNGY	678	
QY	419	GTDRCVWLRLDTYTFWSEFIGPATLIMLVNFIGLIALYKMFHRTAALKPESGCLDNIKSW	478	
DB	679	GRPRYCWMLDFQGFPLMSFLGPVTFILCNNAVIFVTWVKLQKSEINPDMKKLKARAL	738	
QY	479	VIGALAILLCILGLTWARBSIMTINESTYVMAVLFTIFNSLOGMFIFIFICYQAKVRRYRG	538	
DB	739	TITFAIAOLFLLGCTWVEGLFIFDDRSILVLTFTVIFLNLQGAFLYLLHCLLNKKVREBYR	798	
QY	539	K--CLRTHCGSGK-----STESSIGSGKTSGSR	564	
DB	799	KWACL---VAGGSKYSEFTSTTSGTGHNQTRALR	829	
RESULT 2				
EMRL_MOUSE				
ID	EMRL_MOUSE	STANDARD;	PRT;	931 AA.
AC	Q61349;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Cell surface glycoprotein EMRL precursor (EMRL hormone receptor)			
DE	(cell surface glycoprotein P4/80).			
OS	Mus musculus (Mouse).			

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE=Peritoneal cavity;
 RX MEDLINE=96132946; PubMed=8550607;
 RA McKnight A.J., Macfarlane A.J., Dri P., Turley L., Willis A.C.,
 RA Gordon S.;
 RT "Molecular cloning of F4/80, a murine macrophage-restricted cell
 RT surface glycoprotein with homology to the G-protein-linked
 RT transmembrane 7 hormone receptor family";
 RL J. Biol. Chem. 271:486-489(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97312684; PubMed=9169125;
 RA Lin H.H., Stubbs L.J., Mucenski M.L.;
 RT "Identification and characterization of a seven transmembrane hormone
 RT receptor using differential display.";
 RL Genomics 41:301-308(1997).
 CC -!- FUNCTION: PROBABLY INVOLVED IN CELL ADHESION WITHIN TISSUES
 CC AND RECEPTOR SIGNALING.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: IN MACROPHAGES; BUT ABSENT FROM THOSE WHICH
 CC ARE LOCALIZED WITHIN T-CELL AREAS OF LYMPH NODES AND SPLEEN.
 CC LOW LEVEL OF EXPRESSION ON BLOOD MONOCYTES.
 CC -!- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -!- SIMILARITY: CONTAINS 1 GPS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X93328; CAA63720.1; -
 DR EMBL: U66888; AAC53184.1; -
 DR HSP: P35555; IEMN.
 DR MGD: MGI:106912; Emr1.
 DR InterPro: IPR000152; Asx_hydroxyl.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF Ca.
 DR InterPro: IPR000832; GPCR_secretin.
 DR InterPro: IPR000203; PKD_cys-rich.
 DR Pfam: PF00002; 7tm_2; 1.
 DR Pfam: PF00008; EGF; 6.
 DR Pfam: PF01825; GPS; 1.
 DR SMART: SM00179; EGF_CA; 6.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00303; GPS; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 6.
 DR PROSITE: PS01186; EGF_2; 1.
 DR PROSITE: PS01187; EGF_CA; 5.
 DR PROSITE: PS50221; GPS; 1.
 DR PROSITE: PS00650; G_PROTEIN_RECP_F2_2; 1.
 DR PROSITE: PS50261; G_PROTEIN_RECP_F2_4; 1.
 DR G-protein coupled receptor; Transmembrane; Receptor; Glycoprotein;
 KW EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 1 931
 FT DOMAIN 28 644
 FT TRANSMEM 645 672
 FT DOMAIN 673 679
 FT TRANSMEM 680 701
 FT DOMAIN 702 711
 FT TRANSMEM 712 735
 FT DOMAIN 736 754
 FT TRANSMEM 755 776
 FT DOMAIN 777 792
 FT TRANSMEM 793 821
 FT TRANSMEM POTENTIAL.

FT DOMAIN 822 839
 FT TRANSMEM 840 859
 FT DOMAIN 860 874
 FT TRANSMEM 875 897
 FT DOMAIN 898 931
 FT DOMAIN 932 80
 FT DOMAIN 131 132
 FT DOMAIN 133 172
 FT DOMAIN 173 221
 FT DOMAIN 222 271
 FT DOMAIN 272 318
 FT DOMAIN 319 367
 FT DOMAIN 367 641
 FT SITE 506 508
 FT DISULFID 36 48
 FT DISULFID 42 57
 FT DISULFID 59 79
 FT DISULFID 85 98
 FT DISULFID 92 107
 FT DISULFID 109 131
 FT DISULFID 137 149
 FT DISULFID 143 158
 FT DISULFID 160 171
 FT DISULFID 177 189
 FT DISULFID 183 198
 FT DISULFID 200 229
 FT DISULFID 226 230
 FT DISULFID 233 248
 FT DISULFID 250 270
 FT DISULFID 276 286
 FT DISULFID 295 297
 FT DISULFID 297 317
 FT DISULFID 323 336
 FT DISULFID 330 345
 FT DISULFID 347 366
 FT CARBOHYD 148 148
 FT CARBOHYD 167 167
 FT CARBOHYD 229 229
 FT CARBOHYD 259 259
 FT CARBOHYD 263 283
 FT CARBOHYD 405 405
 FT CARBOHYD 417 417
 FT CARBOHYD 474 474
 FT CARBOHYD 498 498
 FT CARBOHYD 706 706
 SQ SEQUENCE 931 AA; 102129 MW; 52963A667E8B76B5 CRC64;
 Query Match 16.4%; Score 749.5; DB 1; Length 931;
 Best Local Similarity 31.9%; Pred. No. 4,7e-41;
 Matches 176; Conservative 101; Mismatches 230; Indels 45; Gaps 11;
 QY 26 LLDVLRNLTP-PEGKDSAAKSLNKANVEYVNNILQPOLANMRDLTTSQRLRAATYMLHT 84
 DB 410 IIDNTCKNSAPVSLQSAATSV-----SLVLRQA-TTWELSKESTSTGLTILLET 459
 QY 85 VEEASVFLVLDNLKTDIVENTDNIKLEVARLSTEGNLDLKRPEN-----MGHGSTIOL 139
 DB 460 VEST-----MVALILITSGNASQMIQTEYLDIESKVINDEKKNESINLAARGKMMV 512
 QY 140 SANTLQNGRNGEIRVAFYLYNNLGYLSTENASMKLTGEALSTNSVIVNSPVITAAIN 199
 DB 513 GCFTIKESYSTGAPVAFSAFAMESVULNERFEDQSPFKLMNSRYV--GGTVGEKK 570
 QY 200 KEFSNKNVLYADPVFTVFKHKQSEENFNRCSSWTSKRTMGYSWGCGRLITTKKTR 259
 DB 571 EDPSK-----PIYTLQHIQPKQKSERPICVSWND--VEDGRMTPSCCEIYEASETH 622
 QY 260 TGCSCNHLTPNAVMAVEYKSHDAVDLDDVTWGGILSLVCLLIGFTFCFFGLQS 319
 DB 623 VCSCKMANLAIIMASGEL-----TWEFSLYIISHGVYISLWCLALATITFLCAVON 677
 QY 320 DRNTIHKNICISLFAVELFLIGINRTPDIACAVFAALHFFFLAFTMPLEGVOLYI 379
 DB 320 DRNTIHKNICISLFAVELFLIGINRTPDIACAVFAALHFFFLAFTMPLEGVOLYI 379

```

Db 678 HNTYMHLCVCEFLAKILFLTGIDKTDNOTACAIIGFLHYLEAFEMVLVEAVMLFL 737
Qy 380 M-----LVEVESESHSRKRYFYLVGCPALITAVSAVDYRSYGTDKVCLRLDTFEIY 434
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 738 MVRRLKAVNVFSSRNIMHLCAGFYGLPVLYIISVOPRGIMNRCNLNRETGFIV 797
Qy 435 SFICPAFLIIMLVNIFGLAIYKMFHHTAIKPESSCLDNKISVIGALAILLGLTWA 494
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 798 SFGLPVCMTITINSVLLAWLTMLVLRKLCVSSEVSKIKDRLTLTFKAIQIFILGCSW 857
Qy 495 FGLMYINESYIMAYLTETINSLOGMFIIFHCVLQKKVREYKCL--RPHCSGKSTE 552
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 858 LGFIQIGPLASIMAYLTETINSLOGAFILHCLLNROVDEYKLLTRKTDLSSHOTS 917
Qy 553 SSISGKTSGSR 564
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 918 GILLSSMPSTSK 929

RESULT 3
EMRL_HUMAN STANDARD; PRT; 886 AA.
AC 014246:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell surface glycoprotein EMRL precursor (EMRL hormone receptor).
GN EMRL.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95324926; PubMed=7601460;
RA Bud V., Chissee S.L., Viegas-Pequignot E., Diriong S., Nguyen V.C.,
RA Roe B.A., Lipinski W.;
RT "EMRL, an unusual member in the family of hormone receptors with
RT seven transmembrane segments." ;
RL Genomics 26:334-344(1995).
CC -1- FUNCTION: PROBABLY INVOLVED IN CELLULAR RESPONSE TO A HORMONE OR
CC AN INTERACTION WITH A PROTEIN LIGAND.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: WIDE EXPRESSION; INCREASED LEVELS IN
CC PERIPHERAL BLOOD MONONUCLEAR CELLS.
CC -1- PTM: N- AND O-GLYCOSYLATED; (POSSIBLE).
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X81479; CAAS7232.1; -
DR HSSP: P00736; IAPQ.
DR Genew; HGNC:3336; EMRL.
DR MIM: 600493; -
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-like.
DR InterPro: IPR000832; GPCR_secretin.
DR InterPro: IPR000203; PKD_cys_rich.
DR Pfam: PF00002; 7tm_2; 1.
DR Pfam: PF01825; GFS; 1.
DR SMART: SM00179; EGF_CA; 5.
DR SMART: SM00001; EGF_like; 1.
DR SMART: SM00303; GPS; 1.

```

```

DR PROSITE: PS00010; ASX_HYDROXYL; 6.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 5.
DR PROSITE: PS50221; GPS; 1.
DR PROSITE: PS00650; G-PROTEIN_RECP_F2_2; 1.
DR PROSITE: PS50261; G-PROTEIN_RECP_F2_4; 1.
KW G-protein coupled receptor; Transmembrane; Receptor; Glycoprotein;
KW EGF-like domain; Repeat; signal.
FT CHAIN 1
FT SIGNAL 17
FT CHAIN 18
FT DOMAIN 18 599
FT TRANSMEM 600 627
FT TRANSMEM 628 634
FT TRANSMEM 635 656
FT TRANSMEM 657 666
FT TRANSMEM 667 690
FT TRANSMEM 691 709
FT TRANSMEM 710 731
FT DOMAIN 732 747
FT TRANSMEM 748 776
FT DOMAIN 777 794
FT TRANSMEM 795 814
FT DOMAIN 815 829
FT TRANSMEM 830 852
FT DOMAIN 853 866
FT DOMAIN 867 886
FT DOMAIN 887 917
FT DOMAIN 918 948
FT DOMAIN 949 979
FT DOMAIN 980 1010
FT DOMAIN 1011 1041
FT DOMAIN 1042 1072
FT DOMAIN 1073 1103
FT DOMAIN 1104 1134
FT DOMAIN 1135 1165
FT DOMAIN 1166 1196
FT DOMAIN 1197 1227
FT DOMAIN 1228 1258
FT DOMAIN 1259 1289
FT DOMAIN 1290 1320
FT DOMAIN 1321 1351
FT DOMAIN 1352 1382
FT DOMAIN 1383 1413
FT DOMAIN 1414 1444
FT DOMAIN 1445 1475
FT DOMAIN 1476 1506
FT DOMAIN 1507 1537
FT DOMAIN 1538 1568
FT DOMAIN 1569 1599
FT DOMAIN 1600 1630
FT DOMAIN 1631 1661
FT DOMAIN 1662 1692
FT DOMAIN 1693 1723
FT DOMAIN 1724 1754
FT DOMAIN 1755 1785
FT DOMAIN 1786 1816
FT DOMAIN 1817 1847
FT DOMAIN 1848 1878
FT DOMAIN 1879 1909
FT DOMAIN 1910 1940
FT DOMAIN 1941 1971
FT DOMAIN 1972 2002
FT DOMAIN 2003 2033
FT DOMAIN 2034 2064
FT DOMAIN 2065 2095
FT DOMAIN 2096 2126
FT DOMAIN 2127 2157
FT DOMAIN 2158 2188
FT DOMAIN 2189 2219
FT DOMAIN 2220 2250
FT DOMAIN 2251 2281
FT DOMAIN 2282 2312
FT DOMAIN 2313 2343
FT DOMAIN 2344 2374
FT DOMAIN 2375 2405
FT DOMAIN 2406 2436
FT DOMAIN 2437 2467
FT DOMAIN 2468 2498
FT DOMAIN 2499 2529
FT DOMAIN 2530 2560
FT DOMAIN 2561 2591
FT DOMAIN 2592 2622
FT DOMAIN 2623 2653
FT DOMAIN 2654 2684
FT DOMAIN 2685 2715
FT DOMAIN 2716 2746
FT DOMAIN 2747 2777
FT DOMAIN 2778 2808
FT DOMAIN 2809 2839
FT DOMAIN 2840 2870
FT DOMAIN 2871 2901
FT DOMAIN 2902 2932
FT DOMAIN 2933 2963
FT DOMAIN 2964 2994
FT DOMAIN 2995 3025
FT DOMAIN 3026 3056
FT DOMAIN 3057 3087
FT DOMAIN 3088 3118
FT DOMAIN 3119 3149
FT DOMAIN 3150 3180
FT DOMAIN 3181 3211
FT DOMAIN 3212 3242
FT DOMAIN 3243 3273
FT DOMAIN 3274 3304
FT DOMAIN 3305 3335
FT DOMAIN 3336 3366
FT DOMAIN 3367 3397
FT DOMAIN 3398 3428
FT DOMAIN 3429 3459
FT DOMAIN 3460 3490
FT DOMAIN 3491 3521
FT DOMAIN 3522 3552
FT DOMAIN 3553 3583
FT DOMAIN 3584 3614
FT DOMAIN 3615 3645
FT DOMAIN 3646 3676
FT DOMAIN 3677 3707
FT DOMAIN 3708 3738
FT DOMAIN 3739 3769
FT DOMAIN 3770 3800
FT DOMAIN 3801 3831
FT DOMAIN 3832 3862
FT DOMAIN 3863 3893
FT DOMAIN 3894 3924
FT DOMAIN 3925 3955
FT DOMAIN 3956 3986
FT DOMAIN 3987 4017
FT DOMAIN 4018 4048
FT DOMAIN 4049 4079
FT DOMAIN 4080 4110
FT DOMAIN 4111 4141
FT DOMAIN 4142 4172
FT DOMAIN 4173 4203
FT DOMAIN 4204 4234
FT DOMAIN 4235 4265
FT DOMAIN 4266 4296
FT DOMAIN 4297 4327
FT DOMAIN 4328 4358
FT DOMAIN 4359 4389
FT DOMAIN 4390 4420
FT DOMAIN 4421 4451
FT DOMAIN 4452 4482
FT DOMAIN 4483 4513
FT DOMAIN 4514 4544
FT DOMAIN 4545 4575
FT DOMAIN 4576 4606
FT DOMAIN 4607 4637
FT DOMAIN 4638 4668
FT DOMAIN 4669 4699
FT DOMAIN 4700 4730
FT DOMAIN 4731 4761
FT DOMAIN 4762 4792
FT DOMAIN 4793 4823
FT DOMAIN 4824 4854
FT DOMAIN 4855 4885
FT DOMAIN 4886 4916
FT DOMAIN 4917 4947
FT DOMAIN 4948 4978
FT DOMAIN 4979 5009
FT DOMAIN 5010 5040
FT DOMAIN 5041 5071
FT DOMAIN 5072 5102
FT DOMAIN 5103 5133
FT DOMAIN 5134 5164
FT DOMAIN 5165 5195
FT DOMAIN 5196 5226
FT DOMAIN 5227 5257
FT DOMAIN 5258 5288
FT DOMAIN 5289 5319
FT DOMAIN 5320 5350
FT DOMAIN 5351 5381
FT DOMAIN 5382 5412
FT DOMAIN 5413 5443
FT DOMAIN 5444 5474
FT DOMAIN 5475 5505
FT DOMAIN 5506 5536
FT DOMAIN 5537 5567
FT DOMAIN 5568 5598
FT DOMAIN 5599 5629
FT DOMAIN 5630 5660
FT DOMAIN 5661 5691
FT DOMAIN 5692 5722
FT DOMAIN 5723 5753
FT DOMAIN 5754 5784
FT DOMAIN 5785 5815
FT DOMAIN 5816 5846
FT DOMAIN 5847 5877
FT DOMAIN 5878 5908
FT DOMAIN 5909 5939
FT DOMAIN 5940 5970
FT DOMAIN 5971 6001
FT DOMAIN 6002 6032
FT DOMAIN 6033 6063
FT DOMAIN 6064 6094
FT DOMAIN 6095 6125
FT DOMAIN 6126 6156
FT DOMAIN 6157 6187
FT DOMAIN 6188 6218
FT DOMAIN 6219 6249
FT DOMAIN 6250 6280
FT DOMAIN 6281 6311
FT DOMAIN 6312 6342
FT DOMAIN 6343 6373
FT DOMAIN 6374 6404
FT DOMAIN 6405 6435
FT DOMAIN 6436 6466
FT DOMAIN 6467 6496
FT DOMAIN 6497 6527
FT DOMAIN 6528 6558
FT DOMAIN 6559 6589
FT DOMAIN 6590 6620
FT DOMAIN 6621 6651
FT DOMAIN 6652 6682
FT DOMAIN 6683 6713
FT DOMAIN 6714 6744
FT DOMAIN 6745 6775
FT DOMAIN 6776 6806
FT DOMAIN 6807 6837
FT DOMAIN 6838 6868
FT DOMAIN 6869 6899
FT DOMAIN 6900 6930
FT DOMAIN 6931 6961
FT DOMAIN 6962 6992
FT DOMAIN 6993 7023
FT DOMAIN 7024 7054
FT DOMAIN 7055 7085
FT DOMAIN 7086 7116
FT DOMAIN 7117 7147
FT DOMAIN 7148 7178
FT DOMAIN 7179 7209
FT DOMAIN 7210 7240
FT DOMAIN 7241 7271
FT DOMAIN 7272 7302
FT DOMAIN 7303 7333
FT DOMAIN 7334 7364
FT DOMAIN 7365 7395
FT DOMAIN 7396 7426
FT DOMAIN 7427 7457
FT DOMAIN 7458 7488
FT DOMAIN 7489 7519
FT DOMAIN 7520 7550
FT DOMAIN 7551 7581
FT DOMAIN 7582 7612
FT DOMAIN 7613 7643
FT DOMAIN 7644 7674
FT DOMAIN 7675 7705
FT DOMAIN 7706 7736
FT DOMAIN 7737 7767
FT DOMAIN 7768 7798
FT DOMAIN 7799 7829
FT DOMAIN 7830 7860
FT DOMAIN 7861 7891
FT DOMAIN 7892 7922
FT DOMAIN 7923 7953
FT DOMAIN 7954 7984
FT DOMAIN 7985 8015
FT DOMAIN 8016 8046
FT DOMAIN 8047 8077
FT DOMAIN 8078 8108
FT DOMAIN 8109 8139
FT DOMAIN 8140 8170
FT DOMAIN 8171 8201
FT DOMAIN 8202 8232
FT DOMAIN 8233 8263
FT DOMAIN 8264 8294
FT DOMAIN 8295 8325
FT DOMAIN 8326 8356
FT DOMAIN 8357 8387
FT DOMAIN 8388 8418
FT DOMAIN 8419 8449
FT DOMAIN 8450 8480
FT DOMAIN 8481 8511
FT DOMAIN 8512 8542
FT DOMAIN 8543 8573
FT DOMAIN 8574 8604
FT DOMAIN 8605 8635
FT DOMAIN 8636 8666
FT DOMAIN 8667 8697
FT DOMAIN 8698 8728
FT DOMAIN 8729 8759
FT DOMAIN 8760 8790
FT DOMAIN 8791 8821
FT DOMAIN 8822 8852
FT DOMAIN 8853 8883
FT DOMAIN 8884 8914
FT DOMAIN 8915 8945
FT DOMAIN 8946 8976
FT DOMAIN 8977 9007
FT DOMAIN 9008 9038
FT DOMAIN 9039 9069
FT DOMAIN 9070 9100
FT DOMAIN 9101 9131
FT DOMAIN 9132 9162
FT DOMAIN 9163 9193
FT DOMAIN 9194 9224
FT DOMAIN 9225 9255
FT DOMAIN 9256 9286
FT DOMAIN 9287 9317
FT DOMAIN 9318 9348
FT DOMAIN 9349 9379
FT DOMAIN 9380 9410
FT DOMAIN 9411 9441
FT DOMAIN 9442 9472
FT DOMAIN 9473 9503
FT DOMAIN 9504 9534
FT DOMAIN 9535 9565
FT DOMAIN 9566 9596
FT DOMAIN 9597 9627
FT DOMAIN 9628 9658
FT DOMAIN 9659 9689
FT DOMAIN 9690 9720
FT DOMAIN 9721 9751
FT DOMAIN 9752 9782
FT DOMAIN 9783 9813
FT DOMAIN 9814 9844
FT DOMAIN 9845 9875
FT DOMAIN 9876 9906
FT DOMAIN 9907 9937
FT DOMAIN 9938 9968
FT DOMAIN 9969 9999
FT DOMAIN 10000 10030
FT DOMAIN 10031 10061
FT DOMAIN 10062 10092
FT DOMAIN 10093 10123
FT DOMAIN 10124 10154
FT DOMAIN 10155 10185
FT DOMAIN 10186 10216
FT DOMAIN 10217 10247
FT DOMAIN 10248 10278
FT DOMAIN 10279 10309
FT DOMAIN 10310 10340
FT DOMAIN 10341 10371
FT DOMAIN 10372 10402
FT DOMAIN 10403 10433
FT DOMAIN 10434 10464
FT DOMAIN 10465 10495
FT DOMAIN 10496 10526
FT DOMAIN 10527 10557
FT DOMAIN 10558 10588
FT DOMAIN 10589 10619
FT DOMAIN 10620 10650
FT DOMAIN 10651 10681
FT DOMAIN 10682 10712
FT DOMAIN 10713 10743
FT DOMAIN 10744 10774
FT DOMAIN 10775 10805
FT DOMAIN 10806 10836
FT DOMAIN 10837 10867
FT DOMAIN 10868 10898
FT DOMAIN 10899 10929
FT DOMAIN 10930 10960
FT DOMAIN 10961 10991
FT DOMAIN 10992 11022
FT DOMAIN 11023 11053
FT DOMAIN 11054 11084
FT DOMAIN 11085 11115
FT DOMAIN 11116 11146
FT DOMAIN 11147 11177
FT DOMAIN 11178 11208
FT DOMAIN 11209 11239
FT DOMAIN 11240 11270
FT DOMAIN 11271 11301
FT DOMAIN 11302 11332
FT DOMAIN 11333 11363
FT DOMAIN 11364 11394
FT DOMAIN 11395 11425
FT DOMAIN 11426 11456
FT DOMAIN 11457 11487
FT DOMAIN 11488 11518
FT DOMAIN 11519 11549
FT DOMAIN 11550 11580
FT DOMAIN 11581 11611
FT DOMAIN 11612 11642
FT DOMAIN 11643 11673
FT DOMAIN 11674 11704
FT DOMAIN 11705 11735
FT DOMAIN 11736 11766
FT DOMAIN 11767 11797
FT DOMAIN 11798 11828
FT DOMAIN 11829 11859
FT DOMAIN 11860 11890
FT DOMAIN 11891 11921
FT DOMAIN 11922 11952
FT DOMAIN 11953 11983
FT DOMAIN 11984 12014
FT DOMAIN 12015 12045
FT DOMAIN 12046 12076
FT DOMAIN 12077 12107
FT DOMAIN 12108 12138
FT DOMAIN 12139 12169
FT DOMAIN 12170 12200
FT DOMAIN 12201 12231
FT DOMAIN 12232 12262
FT DOMAIN 12263 12293
FT DOMAIN 12294 12324
FT DOMAIN 12325 12355
FT DOMAIN 12356 12386
FT DOMAIN 12387 12417
FT DOMAIN 12418 12448
FT DOMAIN 12449 12479
FT DOMAIN 12480 12510
FT DOMAIN 12511 12541
FT DOMAIN 12542 12572
FT DOMAIN 12573 12603
FT DOMAIN 12604 12634
FT DOMAIN 12635 12665
FT DOMAIN 12666 12696
FT DOMAIN 12697 12727
FT DOMAIN 12728 12758
FT DOMAIN 12759 12789
FT DOMAIN 12790 12820
FT DOMAIN 12821 12851
FT DOMAIN 12852 12882
FT DOMAIN 12883 12913
FT DOMAIN 12914 12944
FT DOMAIN 12945 12975
FT DOMAIN 12976 13006
FT DOMAIN 13007 13037
FT DOMAIN 13038 13068
FT DOMAIN 13069 13100
FT DOMAIN 13101 13131
FT DOMAIN 13132 13162
FT DOMAIN 13163 13193
FT DOMAIN 13194 13224
FT DOMAIN 13225 13255
FT DOMAIN 13256 13286
FT DOMAIN 13287 13317
FT DOMAIN 13318 13348
FT DOMAIN 13349 13380
FT DOMAIN 13381 13411
FT DOMAIN 13412 13442
FT DOMAIN 13443 13473
FT DOMAIN 13474 13504
FT DOMAIN 13505 13535
FT DOMAIN 13536 13566
FT DOMAIN 13567 13597
FT DOMAIN 13598 13628
FT DOMAIN 13629 13659
FT DOMAIN 13660 13690
FT DOMAIN 13691 13721
FT DOMAIN 13722 13752
FT DOMAIN 13753 13783
FT DOMAIN 13784 13814
FT DOMAIN 13815 13845
FT DOMAIN 13846 13876
FT DOMAIN 13877 13907
FT DOMAIN 13908 13938
FT DOMAIN 13939 13969
FT DOMAIN 13970 14000
FT DOMAIN 14001 14031
FT DOMAIN 14032 14062
FT DOMAIN 14063 14093
FT DOMAIN 14094 14124
FT DOMAIN 14125 14155
FT DOMAIN 14156 14186
FT DOMAIN 14187 14217
FT DOMAIN 14218 14248
FT DOMAIN 14249 14279
FT DOMAIN 14280 14310
FT DOMAIN 14311 14341
FT DOMAIN 14342 14372
FT DOMAIN 14373 14403
FT DOMAIN 14404 14434
FT DOMAIN 14435 14465
FT DOMAIN 14466 14496
FT DOMAIN 14497 14527
FT DOMAIN 14528 14558
FT DOMAIN 14559 14589
FT DOMAIN 14590 14620
FT DOMAIN 14621 14651
FT DOMAIN 14652 14682
FT DOMAIN 14683 14713
FT DOMAIN 14714 14744
FT DOMAIN 14745 14775
FT DOMAIN 14776 14806
FT DOMAIN 14807 14837
FT DOMAIN 14838 14868
FT DOMAIN 14869 14899
FT DOMAIN 14900 14930
FT DOMAIN 14931 14961
FT DOMAIN 14962 14992
FT DOMAIN 14993 15023
FT DOMAIN 15024 15054
FT DOMAIN 15055 15085
FT DOMAIN 15086 15116
FT DOMAIN 15117 15147
FT DOMAIN 15148 15178
FT DOMAIN 15179 15209
FT DOMAIN 15210 15240
FT DOMAIN 15241 15271
FT DOMAIN 15272 15302
FT DOMAIN 15303 15333
FT DOMAIN 15334 15364
FT DOMAIN 15365 15395
FT DOMAIN 15396 15426
FT DOMAIN 15427 15457
FT DOMAIN 15458 15488
FT DOMAIN 15489 15519
FT DOMAIN 15520 15550
FT DOMAIN 15551 15581
FT DOMAIN 15582 15612
FT DOMAIN 15613 15643
FT DOMAIN 15644 15674
FT DOMAIN 15675 15705
FT DOMAIN 15706 15736
FT DOMAIN 15737 15767
FT DOMAIN 15768 15798
FT DOMAIN 15799 15829
FT DOMAIN 15830 15860
FT DOMAIN 15861 15891
FT DOMAIN 15892 15922
FT DOMAIN 15923 15953
FT DOMAIN 15954 15984
FT DOMAIN 15985 16015
FT DOMAIN 16016 16046
FT DOMAIN 16047 16077
FT DOMAIN 16078 16108
FT DOMAIN 16109 16139
FT DOMAIN 16140 16170
FT DOMAIN 16171 16201
FT DOMAIN 16202 16232
FT DOMAIN 16233 16263
FT DOMAIN 16264 16294
FT DOMAIN 16295 16325
FT DOMAIN 16326 16356
FT DOMAIN 16357 16387
FT DOMAIN 16388 16418
FT DOMAIN 16419 16449
FT DOMAIN 16450 16480
FT DOMAIN 16481 16511
FT DOMAIN 16512 16542
FT DOMAIN 16543 16573
FT DOMAIN 16574 16604
FT DOMAIN 16605 16635
FT DOMAIN 16636 16666
FT DOMAIN 16667 16697
FT DOMAIN 16698 16728
FT DOMAIN 16729 16759
FT DOMAIN 16760 16790
FT DOMAIN 16791 16821
FT DOMAIN 16822 16852
FT DOMAIN 16853 16883
FT DOMAIN 16884 16914
FT DOMAIN 16915 16945
FT DOMAIN 16946 16976
FT DOMAIN 16977 17007
FT DOMAIN 17008 17038
FT DOMAIN 17039 17069
FT DOMAIN 17070 17100
FT DOMAIN 17101 17131
FT DOMAIN 17132 17162
FT DOMAIN 17163 17193
FT DOMAIN 17194 17224
FT DOMAIN 17225 17255
FT DOMAIN 17256 17286
FT DOMAIN 17287 17317
FT DOMAIN 17318 17348
FT DOMAIN 17349 17380
FT DOMAIN 17381 17411
FT DOMAIN 17412 17442
FT DOMAIN 17443 17473
FT DOMAIN 17474 17504
FT DOMAIN 17505 17535
FT DOMAIN 17536 17566
FT DOMAIN 17567 17597
FT DOMAIN 17598 17628
FT DOMAIN 17629 17659
FT DOMAIN 17660 17690
FT DOMAIN 17691 17721
FT DOMAIN 17722 17752
FT DOMAIN 17753 17783
FT DOMAIN 17784 17814
FT DOMAIN 17815 17845
FT DOMAIN 17846 17876
FT DOMAIN 17877 17907
FT DOMAIN 17908 17938
FT DOMAIN 17939 17969
FT DOMAIN 17970 18000
FT DOMAIN 18001 18031
FT DOMAIN 18032 18062
FT DOMAIN 18063 18093
FT DOMAIN 18094 18124
FT DOMAIN 18125 18155
FT DOMAIN 18156 18186
FT DOMAIN 18187 18217
FT DOMAIN 18218 18248
FT DOMAIN 18249 18279
FT DOMAIN 18280 18310
FT DOMAIN 18311 18341
FT DOMAIN 18342 18372
FT DOMAIN 18373 18403
FT DOMAIN 18404 18434
FT DOMAIN 18435 18465
FT DOMAIN 18466 18496
FT DOMAIN 18497 18527
FT DOMAIN 18528 18558
FT DOMAIN 18559 18589
FT DOMAIN 18590 18620
FT DOMAIN 18621 18651
FT DOMAIN 18652 18682
FT DOMAIN 18683 18713
FT DOMAIN 18714 18744
FT DOMAIN 18745 18775
FT DOMAIN 18776 18806
FT DOMAIN 18807 18837
FT DOMAIN 18838 18868
FT DOMAIN 18869 18900
FT DOMAIN 18901 18931
FT DOMAIN 18932 18962
FT DOMAIN 18963 18993
FT DOMAIN 18994 19024
FT DOMAIN 19025 19055
FT DOMAIN 19056 19086
FT DOMAIN 19087 19117
FT DOMAIN 19118 19148
FT DOMAIN 19149 19179
FT DOMAIN 19180 19210
FT DOMAIN 19211 19241
FT DOMAIN 19242 19272
FT DOMAIN 19273 19303
FT DOMAIN 19304 19334
FT DOMAIN 19335 19365
FT DOMAIN 19366 19396
FT DOMAIN 19397 19427
FT DOMAIN 19428 19458
FT DOMAIN 19459 19489
FT DOMAIN 19490 19520
FT DOMAIN 19521 19551
FT DOMAIN 19552 19582
FT DOMAIN 19583 19613
FT DOMAIN 19614 19644
FT DOMAIN 19645 19675
FT DOMAIN 19676 19706
FT DOMAIN 19707 19737
FT DOMAIN 19738 19768
FT DOMAIN 19769 19800
FT DOMAIN 19801 19831
FT DOMAIN 19832 19862
FT DOMAIN 19863 19893
FT DOMAIN 19894 19924
FT DOMAIN 19925 19955
FT DOMAIN 19956 19986
FT DOMAIN 19987 20017
FT DOMAIN 20018 20048
FT DOMAIN 20049 20079
FT DOMAIN 20080 20110
FT DOMAIN 20111 20141
FT DOMAIN 20142 20172
FT DOMAIN 20173 20203
FT DOMAIN 20204 20234
FT DOMAIN 20235 20265
FT DOMAIN 20266 20296
FT DOMAIN 20297 20327
FT DOMAIN 20328 20358
FT DOMAIN 20359 20389
FT DOMAIN 20390 20420
FT DOMAIN 20421 20451
FT DOMAIN 20452 20482
FT DOMAIN 20483 20513
FT DOMAIN 20514 20544
FT DOMAIN 20545 20575
FT DOMAIN 20576 20606
FT DOMAIN 20607 20637
FT DOMAIN 20638 20668
FT DOMAIN 20669 20700
FT DOMAIN 20701 20731
FT DOMAIN 20732 20762
FT DOMAIN 20763 20793
FT DOMAIN 20794 20824
FT DOMAIN 20825 20855
FT DOMAIN 20856 20886
FT DOMAIN 20887 20917
FT DOMAIN 20918 20948
FT DOMAIN 20949 20979
FT DOMAIN 20980 21010
FT DOMAIN 21011 21041
FT DOMAIN 21042 21072
FT DOMAIN 21073 21103
FT DOMAIN 21104 21134
FT DOMAIN 21135 21165
FT DOMAIN 21166 21196
FT DOMAIN 21197 21227
FT DOMAIN 21228 21258
FT DOMAIN 21259 21289
FT DOMAIN 21290 21320
FT DOMAIN 21321 21351
FT DOMAIN 21352 21382
FT DOMAIN 21383 21413
FT DOMAIN 21414 21444
FT DOMAIN 21445 21475
FT DOMAIN 21476 21506
FT DOMAIN 21507 21537
FT DOMAIN 21538 21568
FT DOMAIN 21569 21599
FT DOMAIN 21600 21630
FT DOMAIN 21631 21661
FT DOMAIN 21662 21692
FT DOMAIN 21693 21723
FT DOMAIN 21724 21754
FT DOMAIN 21755 21785
FT DOMAIN 21786 21816
FT DOMAIN 21817 21847
FT DOMAIN 21848 21878
FT DOMAIN 21879 21909
FT DOMAIN 21910 21940
FT DOMAIN 21941 21971
FT DOMAIN 21972 22002
FT DOMAIN 22003 22033
FT DOMAIN 22034 22064
FT DOMAIN 22065 22095
FT DOMAIN 22096 22126
FT DOMAIN 22127 22157
FT DOMAIN 22158 22188
FT DOMAIN 22189 22219
FT DOMAIN 22220 22250
FT DOMAIN 22251 22281
FT DOMAIN 22282 22312
FT DOMAIN 22313 22343
FT DOMAIN 22344 22374
FT DOMAIN 22375 22405
FT DOMAIN 22406 22436
FT DOMAIN 22437 22467
FT DOMAIN 22468 22498
FT DOMAIN 22499 22529
FT DOMAIN 22530 22560
FT DOMAIN 22561 22591
FT DOMAIN 22592 22622
FT DOMAIN 22623 22653
FT DOMAIN 22654 22684
FT DOMAIN 22685 22715
FT DOMAIN 22716 22746
FT DOMAIN 22747 22777
FT DOMAIN 22778 22808
FT DOMAIN 22809 22839
FT DOMAIN 22840 22870
FT DOMAIN 22871 22901
FT DOMAIN 22902 22932
FT DOMAIN 22933 22963
FT DOMAIN 22964 22994
FT DOMAIN 22995 23025
FT DOMAIN 23026 23056
FT DOMAIN 23057 23087
FT DOMAIN 23088 23118
FT DOMAIN 23119 23149
FT DOMAIN 23150 23180
FT DOMAIN 23181 23211
FT DOMAIN 23212 23242
FT DOMAIN 23243 23273
FT DOMAIN 23274 23304
FT DOMAIN 23305 23335
FT DOMAIN 23336 23366
FT DOMAIN 23367 23397
FT DOMAIN 23398 23428
FT DOMAIN 23429 23459
FT DOMAIN 23460 23490
FT DOMAIN 23491 23521
FT DOMAIN 23522 23552
FT DOMAIN 23553 23583
FT DOMAIN 23584 23614
FT DOMAIN 23615 23645
FT DOMAIN 23646 23676
FT DOMAIN 23677 23707
FT DOMAIN 23708 23738
FT DOMAIN 23739 2376
```


FT	CARBOHYD	793	793	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1129	1129	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1154	1154	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1228	1228	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1264	1264	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1274	1274	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1302	1302	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1591	1591	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1638	1638	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1655	1655	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	1994	1994	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2118	2118	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2137	2137	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2144	2144	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2155	2155	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2160	2160	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2272	2272	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2430	2430	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2452	2452	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	2538	2538	N-LINKED (GLCNAC . .)	(POTENTIAL)
SQ	SEQUENCE	3034 AA:	330477 MM:	EPF83180AF5ED8A8 CRC64;	
Query Match					
Best Local Similarity		24.5%	Score 700.5:	DB 1:	Length 3034;
Matches 237:		Conservative 166:	Mismatches 356:	Indels 207:	Gaps
OY	1 AEOTRNHLNAGDIYV--SVRAMDVLGDLVDVORNLTPGSGKDSAAKSLR--KAMETVYNN 56				
Db	2150 AKALRNMTQGSSTLFNGNDPATAVOLLARI-LQHESHQGGDLATAEAFNHEDVHTSQA 2208				
OY	57 LLQPOLANMRDLTTSQDLRAATMLH-----TVESAFAVLADNLKTDI 101				
Db	2209 LLAPATEASWEQIORS-EGAAADLIRHFEAFSGNVAARYKRYLRFLRFVLTANMILAADI 2267				
OY	102 VRE-----NTNDIKLEVAR-----LSTEGNLEDAKPPENMGHGTIDL----- 139				
Db	2268 FDKLNTFGAOUVRREDIQEELPRELESSVSFPADTRKPREKK-EGGVYVALIKRRTPPLTA 2322				
OY	140 -----SANTLKONGRNGEIRYAF-VLYNNIGPYL---STENASMKLGTEALST 183				
Db	2327 QPREPAREHETSSSKRRRHNPDEQGFALVALVYIRKLGOLLPREHYDRHSLRLPRRP-- 2383				
OY	184 NHSVIVNSPVTTAAINKFEISKUYIADPVVFTYKTKHSGEENENPNCSSWTSKRT-MTG 242				
Db	2384 ----VINTPVVASAVYSEGRPLPSSLDORPLIVEFSLEETDEERSKRPVCYFVWNHSLDTGGG 2439				
OY	243 YMSOGGCRLLTNTHTTCTGSCNHLNTNPAVMAHVEKNSDAVNDLDDITVWGLISLY 302				
Db	2440 GMSAKGCELLSRNTHTCOSSHASCAVALMDSIREHGEV--LPLKITYTAALISLTV 2496				
OY	303 CLLICIFPFCFGRLOQSDRMTIKNLCISLFEVALLFLGINKTDPRIACAFAALLHFF 362				
Db	2497 ALLVAFVLLSLVRLRSRNLNHSIKHNLIAALFFSQLLFNGTENTPRCLCTVVALILHY 2556				
OY	363 FLAFTWMELEGVOLYITMLVFESEHSRRKTYIYLVGCMGPRALIAVSAANDYSTGTDK 422				
Db	2557 SMGFATVTLVENLITVRYMLTEVENIDTPGRMYRFHVGWGIAPILVGLAAGLDPOGGRPD 2616				
OY	423 VCMRLDTPYFMSIRGPATILMLN-VIPL---GIATUMPHNTALRKESCLINISM 478				
Db	2617 FCMVSLDLOTLTSPAGVAGVITVITVYFLSAKVSQCKRNNH---YERKKVVSMLRT- 2671				
OY	479 VIGAIALCLLGLTWAGFLWYINESTVYMALETTINSLOGMFIHCVLOKRYKREY- 537				
Db	2672 ---AFILLLLVTAIWWLGLLAIVSDPLSFHYLFAASCQIGIFVLLFHCVAHREVRKHLR 2722				
OY	538 ----GKCLIRHCSGKSTSGSSIGSGTRPGRYSTGSSQSRIRRMANDYRKOSSESF 593				
Db	2729 AVLAGKKILOLD--DSATTRATLTLLRSINCNT---YSEGR-----DMLR----- 2767				
OY	594 ITGQINSASLNRGLLNARDTSY----MDPLRNGHNGNSTYASGEYLSNCVQIIDR 649				
Db	2768 -TALGESTASYLD-----STTRDGVQKLVSSDPARGANGEDPT-----SFIPRNSK 2814				

```

QY 650 GYNHNETALEKKIKELTSNYIPSYLNNHRSSEQNRLMN-----KLVNLT 696
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2815 ANGPDSDSSELSLDEHSSSVASHTSDSEDDGEAEKDMWNPAGHSTPKADLANIV 2874
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 697 GSGREDDAIVLDATSFNHEESLGLEL-----IHE-----ESPAPLIPRY 737
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2875 PAGWDESLAGSDSELDTEPHILKVKYVELHROAGNHCGRDPSDECVLAKPVAV 2934
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 738 YSTENHNPHTYRRRIPOHSESF-----PLLT--NHTEDLOSPPHDSLY 782
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2935 LSSQ-----PQQRKGIILKNKYTPPLPEQPLKSLRLKRLADCCQSPTSRT 2982
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 783 TSMPTLACVAATESYTT--STOTEPPEAKGDAEDVYKSMPLGSRNHVOLHTYYQLG 840
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2983 SSLSGSDGVHATDCVITITKTRREP-----GREHLNGVAMNVRIG 3022
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 841 RGSSDQ 846
      : : :
DB 3023 SQAANG 3028

RESULT 5
CLRL_HUMAN STANDARD; PRT; 3014 AA.
AC Q9NYO6; Q9Y526; Q9Y506; Q95722; Q9BMO5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Caderin EGF LAG seven-pass G-type receptor 1 precursor (Flamingo
DE homolog 2) (hm12).
GN CELSRL OR CDHP9 OR FM12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20202599; PubMed=10716726;
RA Wu Q., Maniatis T.;
RT "Large exons encoding multiple ectodomains are a characteristic
RL feature of protocadherin genes.";
RN Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiwich R., Beare D.M.,
RA Clamp M., Smitk L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burtil W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conway D., Corby N.R., Coville G.J., Cox A.V., Davys J., Dawson E.,
RA Dhami P.D., Dockree C., Dowsorth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leverisha M.A., Lloyd C., Lloyd D.M.,
RA McLaren I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA McElay J., McLaren S., McMurray A.A., Milne S.A., Morrison B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sultun J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,

```

```

RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerisky P., Rohlfing T.,
RA Schert P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Marsis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saita S., Budarf M.L.,
RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M.I., Dunamski J.P., Peyrard M., Keira D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenreich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tilihan Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [3]
RP SEQUENCE OF 624-3014 FROM N.A. (ISOFORM 2).
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: Receptor that may have an important role in cell/cell
CC signaling during nervous system formation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC produced by alternative splicing.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- SIMILARITY: CONTAINS 9 CADHERIN DOMAINS.
CC -!- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 GPS DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; AF231024; AAF61930.1; -
CC EMBL; AL021392; CAB50707.1; -
CC EMBL; AL031597; CAB45020.1; AL1_INIT.
CC EMBL; AL031588; CAB38413.1; -
CC EMBL; BC000059; AAHD0059.1; -
CC Genew; HGNC:1850; CELSRL.
CC MIM; 604523; -
CC HSSP; P00749; 10RK.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR002126; Cadherin.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000832; GPCR_secretin.
CC InterPro; IPR001879; hormn_receptor.
CC InterPro; IPR002049; laminin_EGF.
CC InterPro; IPR001791; laminin_G.
CC InterPro; IPR000203; PKD_cys-rich.
CC Pfam; PF00002; 7tm_2; 1.
CC Pfam; PF00028; cadherin; 8.
CC Pfam; PF00008; EGF; 6.
CC Pfam; PF01825; GPS; 1.
CC Pfam; PF02793; HRM; 1.
CC Pfam; PF00054; laminin_G; 1.
CC PRINTS; PR00205; CADHERIN.
CC PRINTS; PR00011; EGF_LAMININ.
CC PRINTS; PR00249; GPCRSECRETIN.
CC SMART; SM00112; CA; 9.
CC SMART; SM00180; EGF_Lam; 1.
CC SMART; SM00001; EGF_Like; 6.
CC SMART; SM00303; GPS; 1.
CC SMART; SM00008; Hormn; 1.
CC SMART; SM00282; Lamg; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 2.

```

DR PROSITE; PS00232; CADHERIN_1; 7.
 DR PROSITE; PS00268; CADHERIN_2; 9.
 DR PROSITE; PS00022; EGF_1; 6.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS00221; GFS_1.
 DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; FALSE_NEG.
 DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; FALSE_NEG.
 DR PROSITE; PS00227; G_PROTEIN_RECP_F2_3; 1.
 DR PROSITE; PS00261; G_PROTEIN_RECP_F2_4; 1.
 DR PROSITE; PS01248; LAMININ_TYRE_EGF; 1.
 DR PROSITE; PS00025; LAM_G_DOMAIN; 2.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW EGF-like domain; Calcium-binding; Laminin EGF-like domain; Repeat;
 KW Developmental protein; Hydroxylation; Signal; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 3014
 FT DOMAIN 22 2469 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 2470 2490 1 (POTENTIAL).
 FT DOMAIN 2491 2501 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 2502 2522 2 (POTENTIAL).
 FT DOMAIN 2523 2527 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 2528 2548 3 (POTENTIAL).
 FT DOMAIN 2549 2572 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 2573 2593 4 (POTENTIAL).
 FT DOMAIN 2594 2611 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 2612 2632 5 (POTENTIAL).
 FT DOMAIN 2633 2655 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 2656 2676 6 (POTENTIAL).
 FT DOMAIN 2677 2683 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 2684 2704 7 (POTENTIAL).
 FT DOMAIN 2705 3014 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 3014 353 CADHERIN 1.
 FT DOMAIN 354 459 CADHERIN 2.
 FT TRANSMEM 460 565 CADHERIN 3.
 FT DOMAIN 566 687 CADHERIN 4.
 FT TRANSMEM 688 789 CADHERIN 5.
 FT DOMAIN 790 892 CADHERIN 6.
 FT TRANSMEM 893 999 CADHERIN 7.
 FT DOMAIN 1000 1101 CADHERIN 8.
 FT TRANSMEM 1106 1224 CADHERIN 9.
 FT DOMAIN 1303 1361 EGF-LIKE 1, CALCIUM-BINDING.
 FT TRANSMEM 1363 1399 EGF-LIKE 2, CALCIUM-BINDING.
 FT DOMAIN 1403 1441 EGF-LIKE 3, CALCIUM-BINDING.
 FT TRANSMEM 1442 1646 LAMININ G-LIKE 1.
 FT DOMAIN 1649 1685 LAMININ G-LIKE 4, CALCIUM-BINDING.
 FT TRANSMEM 1689 1870 LAMININ G-LIKE 2.
 FT DOMAIN 1872 1907 EGF-LIKE 5, CALCIUM-BINDING.
 FT TRANSMEM 1908 1946 EGF-LIKE 6, CALCIUM-BINDING.
 FT DOMAIN 1947 1979 EGF-LIKE 7, CALCIUM-BINDING.
 FT TRANSMEM 1981 2016 EGF-LIKE 8, CALCIUM-BINDING.
 FT DOMAIN 2022 2055 LAMININ EGF-LIKE.
 FT TRANSMEM 2408 2460 GPS.
 FT DOMAIN 2659 2663 POLY-LEU.
 FT TRANSMEM 2663 2663 BY SIMILARITY.
 FT DISULFID 1307 1318 BY SIMILARITY.
 FT DISULFID 1312 1349 BY SIMILARITY.
 FT DISULFID 1351 1360 BY SIMILARITY.
 FT DISULFID 1367 1378 BY SIMILARITY.
 FT DISULFID 1372 1387 BY SIMILARITY.
 FT DISULFID 1389 1398 BY SIMILARITY.
 FT DISULFID 1407 1418 BY SIMILARITY.
 FT DISULFID 1412 1428 BY SIMILARITY.
 FT DISULFID 1430 1440 BY SIMILARITY.
 FT DISULFID 1453 1464 BY SIMILARITY.
 FT DISULFID 1658 1673 BY SIMILARITY.
 FT DISULFID 1675 1684 BY SIMILARITY.
 FT DISULFID 1876 1887 BY SIMILARITY.
 FT DISULFID 1881 1896 BY SIMILARITY.
 FT DISULFID 1898 1907 BY SIMILARITY.
 FT DISULFID 1911 1922 BY SIMILARITY.
 FT DISULFID 1916 1934 BY SIMILARITY.
 FT DISULFID 1936 1945 BY SIMILARITY.
 FT DISULFID 1945 1963 BY SIMILARITY.

FT DISULFID 1953 1966 BY SIMILARITY.
 FT DISULFID 1968 1978 BY SIMILARITY.
 FT DISULFID 1985 2000 BY SIMILARITY.
 FT DISULFID 1987 2003 BY SIMILARITY.
 FT DISULFID 2005 2015 BY SIMILARITY.
 Query Match 15.2%; Score 692; DB 1; Length 3014;
 Best local Similarity 24.3%; Pred. No. 1.4e-36;
 Matches 236; Conservative 141; Mismatches 338; Indels 256; Gaps 35;
 QY 17 VRANDVLGLLDVQLRLTLTGCGDSAA---RSLNKAMVEIVNNLQPOLNARDLTSD 73
 DB 2153 VRYAYOLLG---HYLQHSWQOQFDLATODADHEDVHGSALLPATRAAEQIOKSE 2210
 QY 74 OLRAATLTLTVE-----ESAVPLDNLTKTIVRENTNILEVARLS 117
 DB 2211 --GGTAQLRLRLGYSNVANRYRRLRYLRFVLYTANMILAVDIF---DKENFTGARVP 2264
 QY 118 TEGNLEDLKPEPMNGHSTIOLSAN-----TLKONGRN----- 150
 DB 2265 RFDTHIE-EPREL--ESSVSPADFRPEPEKEGPLLBPAGRRTPQTRPGPTGEREA 2321
 QY 151 -----GEIRAFV-LYNNIGPYL-----STENAMKLTETALSNHSYVNSPY 193
 DB 2322 PISRRRRHPDDAQOFAVALYIYRLGQLPBERYDPRRLRL-----PHRPIINTPM 2374
 QY 194 ITAANKPEFNKYLDPV---VETVKHIKSENFNPNCSPMSYSKRT-MTGYSWTOGC 249
 DB 2375 VSTLV---YEGAPLPPELPVLPVEFALLEVERKRPVCVFNNHSLAVGTGWSNRGC 2431
 QY 250 RLTLTKHTTTCSCNHLTNFAVMAHVEKHSADVHDLLDVITWVGILLSTVCLICIF 309
 DB 2432 ELSRNTHVACQCSHTASPAVLM---DISRRENGEVLPIKITYAASVSLAALLVAEV 2488
 QY 310 TPFCEFRLODRMTIHKNCISLFAVELLFLIINRDOYIACAVFALLHFFLAFTW 369
 DB 2489 LLSILVRMLNLSIRKHLVALFELSOLVFIQINDENPFLTVAAILLHYMTSTFAW 2348
 QY 370 MLEGVOLYIMLVEFESEHSRRKRYLVGSGPAPALVAASADVRSYGTDCWMLRLD 429
 DB 2549 TLVESLHVMTLEVRNIDIGPMRFYVVGWGIPLATVGLAVGLDQGYGNPFCMLSLQ 2608
 QY 430 TYFTWSEFGPATLILMLNVFIIGIALYKMFHHPAILKPESGC-----LDNIKSMVIG 481
 DB 2609 DTLIWSFAGIGAVIITNV-----TSVLAKVSCORKNHYGKGIIVSLRT 2656
 QY 482 AIALLCILGLTNAFGLMYINVESTIVIMAYLFTINSLOGMFIIFHCYLOKKVKEKGL 541
 DB 2657 AFLLLILISATWLLGLLAVNRDALSFHYLPAIFSGLGPPVLLFHCVLNDEVRKHLKGLV 2716
 QY 542 ---RTHCCSGKSTES-----SIGSGKTSGS-----RTPGRYSTGSOSRIRRMNDTVRKO 588
 DB 2717 GGRKLHEDSATIRATLITLSLNCNTTFGDPGMLRTDLGESAITSLSIYR--DEGIQID 2774
 QY 589 SESSFITGDNSSASLNREGILNNAARDYSVDTPLPNGNHGSSSIASGEYLSNCOYIID 648
 DB 2775 GVSS-----GL-----VRSGHGEDDA-----SLMP 2794
 QY 649 RGY-----NNHETLLEKKIKELTSLNTPSYLNHHSSEONRLMKLVNNGSGGEDDA 704
 DB 2795 RSCKDPGHDSDSDSELSDEOSS---SYASHSSDSEDD-----GVGAEEK- 2838
 QY 705 IVLDDATSFHHEESLG-----LELHEESDAPLPPRV-----YSTENHOPH 746
 DB 2839 --WDPARGAVHSTPKGDVAVNHYPAGHPDOSLSESDPSGPKRLKVEKYVVELHRE 2896
 QY 747 HYTRRRIPQDHSFPPLLTNHEHTEDLOSPPHSDSLYTMPTLGLAVATESVTTSQTEPP 806
 DB 2897 QGSHR-----GEYPPDQESGARLASSQPP-----DQRKILKNKVTYPP 2937
 QY 807 P-----AKCGAED-----VYKSPNGLSRNHVQLRT 835
 DB 2938 PLTLTEQTLKRLREKLADCEQSPTSRSTSLSGGPGDCAITVKSIPGRERDHLNGVAM 2997


```

FT DISULFID 1149 1161 BY SIMILARITY.
FT DISULFID 1151 1168 BY SIMILARITY.
FT DISULFID 1170 1179 BY SIMILARITY.
FT MOD_RES 816 816 HYDROXYLATION (POTENTIAL).
FT MOD_RES 1035 1035 HYDROXYLATION (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 790 790 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 966 966 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1052 1052 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1123 1123 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1249 1249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1286 1286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1548 1548 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1570 1570 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 2144 AA; 233480 MW; 6EAB89C1BA655ECC CRC64;

```

```

Query Match 14.0%; Score 637.5; DB 1; Length 2144;
Best Local Similarity 26.9%; Pred. No. 3,1e-33;
Matches 180; Conservative 107; Mismatches 258; Indels 123; Gaps 19;

```

```

QY 46 LNKAMETVNN-----LLOPOLNAMDITTSQDLRAATMLHTVEESAFLADNLLKT- 99
DB 1336 VGSALLDAANKRHWEILQOTREGGTAW-----LLOHYEAYASALQNMHRTY 1381
QY 100 -----DIVRENTDNINKLEVALNSTEGNEDLKFP----- 128
DB 1382 LSPFTY---TPNIVISVRLD-KGNPAGTKRPEALRGERPPDETTVILPESVREM 1437
QY 129 -----ENMGHSTIOLASNTLKONGR-----NGEIRVAFVLNNLCOPYSTENASKLTGEA 180
DB 1438 PPMVRSAGPEEAGEETELARQRRHPELSGGEAVASVITIIHTLGLLPHNYDPK---RS 1494
QY 181 LSTNHSTVNSPVITAIANKENFSKNVYLADPVVFTYVKNHKOSEENFNPNCSFMSYSKRTM 240
DB 1495 LRVPKRPIVINTPVVISVHDEDELLPRALDKPVYVQFRLLLETBERTKPICVFNHSTIVS 1554
QY 241 -TGWVSOGORLLTNTKNTHTTSCNHLTNFVLAHVEVHSDAVHLLDVTIVGILL 299
DB 1555 GTGWSMGRGEEVFNRSNHSVSCCNHMTSPAFLM---DVSRRREGELLPKTLTYVALGV 1611
QY 300 SLVCLLCITFCFFRGLOSDRNTIHNKCLISLFEVALLFLIGINTROPDIACAVFALL 359
DB 1612 TLAAIMTFELTLRLALRBNHGIRNLALGALQVLELLGINODLPFACIVAILL 1671
QY 360 HFFFLAAFTWMLLEGVQLYTMLVEVESEHSRRKRYFLVGYGMALIVAAVSAVDYRSYG 419
DB 1672 HFLYLCTFFSWALLEALHLYBALTEVRDVNASPMRFYMLGMVPAFTIGLAVGIDPEGYG 1731
QY 420 TDVVCWRLDTYFIWISIRIGPATLIMLVFLGIALYKMHHAILLKPEGCIDNIR--- 476
DB 1732 NPPOCWLSTIDTLWSAGVAFAVAVSVF-----TL-----TLARSASCAQROGFE 1779
QY 477 -----SWVIGAILLCLLIGTMAFGLMYINESTVIMAYLFTIENSLQGMFIFFHYCLOK 531
DB 1780 KKGPSVLSRSSFTVLLLSLTLALLSVNSDILLFHYLTAACNCVOGPFIFLSVYVLSK 1839
QY 532 KVRKEYGKCLRTHCSCGSKSTESSI--GSGKTSGRPPGRYSTGSQSHIRMMMDTVARK-- 587
DB 1840 EVR-----KALKFACSRKRPSPDPAITKSTLTSSYNCPSPYADG---RLYQPGDSAGSLH 1892
QY 588 -----QSESEFTIGDINSASLN-----REGLLNNAK-----TSVMDTLPLN 625
DB 1893 SASRSGSQSYLIPFLUREESTLNPQGVPPGLDPSGLFMEGQAQOHDPTDSDSLSL 1952
QY 626 GNHGNSTYS 633
DB 1953 DDGSGSYA 1960

```

```

RESULT 7
CLR2_HUMAN
ID CLR2_HUMAN STANDARD; PRT; 2923 AA.
AC 09HC04: 092566;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cadherin EGF LAG seven-pass G-type receptor 2 precursor (Epidermal
DE growth factor-like 2) (Multiple epidermal growth factor-like domains
DE 3) (Flamingo 1).
GN CELSR2 OR CHERF10 OR EGFL2 OR MECF3 OR KIA0279.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20363102; PubMed=10907856;
RA Vincent J.B., Skaug J., Scherer S.W.;
RT "The human homologue of flamingo, EGFL2, encodes a brain-expressed
RT large cadherin-like protein with epidermal growth factor-like domains,
RT and maps to chromosome 1p13.3-p21.1."
RL DNA Res. 7:233-235(2000).
RN [2]
RP SEQUENCE OF 516-2923 FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y.,
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (KIA0201-KIA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain."
RL DNA Res. 3:321-329(1996).
CC -1- FUNCTION: Receptor that may have an important role in cell/cell
CC signalling during nervous system formation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Highest expression in brain and testis.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 9 CADHERIN DOMAINS.
CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL: AF234887; AMG00080.1; -.
DR EMBL: D87469; BAI3407.1; -.
DR HSSP: P15116; INCU.
DR GeneW: HGNC:3231; CELSR2.
DR MIM: 604265; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000832; GPCR_secretin.
DR InterPro: IPR001879; hormn_receptor.
DR InterPro: IPR002049; laminin_Egf.
DR InterPro: IPR001793; laminin_G.
DR InterPro: IPR000203; PKD_cys-rich.
DR Pfam: PF00002; 7tm_2; 1.
DR Pfam: PF00028; cadherin; 8.
DR Pfam: PF00008; EGF; 6.
DR Pfam: PF01825; GPS; 1.
DR Pfam: PF02793; HRK; 1.
DR Pfam: PF00054; laminin_G; 1.

```


Db 2555 KCGVSGSLQPSFAVLLLSAIFLLLSVNSDITLFFHLEATCICIGPFIFLSYVLSK 2614
QY 532 KVRREYKCLRTGCCSGSTRESSI--GSGKTSRGSRTPGRYSTGQSRIRRMMNTVTK-- 587
Db 2615 EVR---KALKIACSRKPSDPDPALTTKSTLSTSYNCPSPIADG---RLYQYGDYSGASLH 2667
QY 588 -----QSESSFTTGIDINSSASLN 605
Db 2668 STSRGSKGPSYIPFLRESALN 2691
RESULT 8
CLR2_MOUSE
ID CLR2_MOUSE STANDARD; PRT; 2920 AA.
AC G9ROM0: Q92R4; Q99K26;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cadherin EGF LAG seven-pass G-type receptor 2 precursor (Flamingo 1)
DE (mfam11).
GN CELSR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=99418630; PubMed=10490098;
RA Usui T., Shima Y., Shimada Y., Hirano S., Burgess R.W., Schwarz T.L.,
RA Takeichi M., Uemura T.;
RT Flamingo, a seven-pass transmembrane cadherin, regulates planar cell
RT polarity under the control of frizzled.;
RL Cell 98:585-595(1999).
RN [2]
RP SEQUENCE OF 1913-2796 FROM N.A. AND TISSUE SPECIFICITY.
RX PubMed=10790539;
RA Formstone C.J., Barclay J., Rees M., Little P.F.R.;
RT "Chromosomal localization of Celstr2 and Celstr3 in the mouse: Celstr3 is
RT a candidate for the floppy (fip) lethal mutant on chromosome 9.";
RL Mamm. Genome 11:392-394(2000).
RN [3]
RP SEQUENCE OF 2014-2920 FROM N.A.
RC TISSUE-Breast tumor;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP DEVELOPMENTAL STAGE.
RX PubMed=11850187;
RA Tissir F., De-Backer O., Goffinet A.M., Lambert de Rouvroit C.A.;
RT "Developmental expression profiles of Celstr (Flamingo) genes in the
RT mouse.";
RL Mech. Dev. 112:157-160(2002).
CC -1- FUNCTION: Receptor that may have an important role in cell/cell
CC signaling during nervous system formation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DEVELOPMENTAL STAGE: Predominantly expressed in the developing
CC CNS, the emerging dorsal root ganglia and cranial ganglia. In the
CC CNS, expression is uniform along the rostrocaudal axis. During
CC gastrulation, it is expressed within the anterior neural ectoderm.
CC At E10, expression is strong in the ventricular zones (VZ) in all
CC sectors of the brain, and lower in the marginal zones (MZ). It is
CC strong in VZ, lower in MZ, except in telecephalic MZ where it is
CC predominant. The intensity is higher in all VZ, and lower in
CC differentiating fields than in VZ, except in the cerebral
CC hemispheres, and to a lesser extent in the tectum and cerebellum.
CC A weak expression is also observed in the fetal lungs, kidney and
CC epithelia. In the newborn and postnatal stages, expression remains
CC restricted to the VZ as well as in migrating and postmigratory
CC cells throughout the brain.
CC -1- TISSUE SPECIFICITY: Expressed in the CNS and in the eye.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.

CC -1- SIMILARITY: CONTAINS 9 CADHERIN DOMAINS.
CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB028499; BAA84070.1; -;
DR EMBL: AF031573; AAC68837.1; -;
DR EMBL: BC005499; AAH05499.1; -;
DR HSP: P00740.1EDM.
DR MGD: MGI:1858235; Celstr2.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR000832; GPCR_secretin.
DR InterPro: IPR001879; hormn_receptor.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR001791; Laminin-G.
DR InterPro: IPR002033; PKC_cys-rich.
DR Pfam: PF00002; 7tm_2; 1.
DR Pfam: PF00028; cadherin; 9.
DR Pfam: PF00008; EGF; 5.
DR Pfam: PF01825; GPS; 1.
DR Pfam: PF02793; HRM; 1.
DR Pfam: PF00054; laminin-G; 2.
DR PRINTS: PR00205; CADHERIN.
DR PRINTS: PR00011; EGF_LAMININ.
DR PRINTS: PR00249; GPCRSECRETIN.
DR SMART: SM00112; CA; 9.
DR SMART: SM00180; EGF_Lam; 1.
DR SMART: SM00001; EGF_Like; 6.
DR SMART: SM00303; GPS_1.
DR SMART: SM00008; Hormr; 1.
DR SMART: SM00282; Lang; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00232; CADHERIN_1; 6.
DR PROSITE: PS0268; CADHERIN_2; 9.
DR PROSITE: PS00022; EGF_1; 6.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS00649; G-PROTEIN_RECEPTOR_F2_1; FALSE_NEG.
DR PROSITE: PS00650; G-PROTEIN_RECEPTOR_F2_2; FALSE_NEG.
DR PROSITE: PS50227; G-PROTEIN_RECEPTOR_F2_3; 1.
DR PROSITE: PS50261; G-PROTEIN_RECEPTOR_F2_4; 1.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 1.
DR PROSITE: PS50221; GPS; 1.
DR PROSITE: PS50025; LAM_G_DOMAIN; 2.
KW G-protein coupled receptor; Transmembrane; glycoprotein;
KW EGF-like domain; Calcium-binding; Laminin EGF-like domain; Repeat;
KW Developmental protein; Hydroxylation; Signal.
FT SIGNAL 1 31
FT CHAIN 32 2920
FT FT
FT DOMAIN 32 2381
FT TRANSSEM 2382 2402
FT DOMAIN 2403 2414
FT TRANSSEM 2415 2434
FT DOMAIN 2435 2439
FT TRANSSEM 2440 2460
FT DOMAIN 2461 2481
FT TRANSSEM 2482 2502
FT DOMAIN 2503 2519
FT TRANSSEM 2520 2540
FT DOMAIN 2541 2564
FT TRANSSEM 2565 2585

Oy	46	LNNAMVEYVANN-----LLOPQALNMMRDLLTTSOOLRAATMLHRTVESAPVLADNLKLT-	99
Db	2112	VGSALLDAANKRMELIQQTEGGTAW-----LLOHYEAYASALAQNMRHTY	2157
Oy	100	---DIVRENTDRIKEVARLSTECNLDLKKP-----	128
Db	2158	LSPFTTY--TPNIIVSYVRLD-KGNFAGTKLPRIEALRGRRPPDVETTYILDESFRREM	2213
Oy	129	----ENMGHSTIOISANTLKONGR-----NGEIRAVFLYNMLGPLYSTFNASMKIGTEA	180
Db	2214	PSMWASAGPEAEQETBELARQRHRRLDELSSQGAAVASVIITHTLAGLLPHNYDDK---RS	2270
Oy	181	LSTNHVYVNSPYTAIAIKERFSNKVVYLADPVYFTWKHKIKOSEBNPNPCSFMSYSKRPM	240
Db	2271	LRPVKRPIVITPAVSISVDDEELLPRALDRPVTVQFRLTEEBERTKPICVFNNHSILVS	2330
Oy	241	-TVYWSPOGRILLTTKTHTCSCMHLNPNVLANHVAKSADVHDLIDVTWTGILL	299
Db	2331	GTOGMSARGEGEVFRNESHVSCQCNHMTSFALM---DMSRREGELTLPLKTLIYVALGV	2387
Oy	300	SIVOLLICIEFECEFGQSDRENTHKNLICLSLFVAELLEFLIGINPTDPIACAFPALL	359
Db	2388	TLAAMLTFELFLLRALKNSHGIRNRULTALAGAOVLFLGINQADLPFACTYAAILL	2447
Oy	360	HFFFLAFTWMLEGOVLYIMLEVPESEHSRKRYFYLVGYGMALIVAAVDYRSG	419
Db	2448	HFVLCFCFSMALDEALHTRALTEYRDVNASMPRFYYMLGWGPAPFTYGLAVGLDEGYG	2507
Oy	420	TBVCWLRLDTYRWISFCIPRALILMLNVIFLGIALYKMFINHTLKRPEGCLDNKK---	476
Db	2508	NPDPCWLSTVYDILTWISFAFPVAFVSMVSF---Ly-----ILSARRSCAAQKGFE	2555
Oy	477	---SWVIGAIALLCGLLGAFLMYINESTVIMAYLFTPFNSLOGMFIFHCVLQK	531
Db	2556	KKGVSGLRSSFFVYLLBLSATMLNALLSVNSDTLLFHFLAACVCYGOPFIIFYLVLEK	2615
Oy	532	KVRKEYGCKLRTHCSGSKSTEISSIGS--GKTSGSFTPGRYSTGSQSIIRRMNDYVRK--	587
Db	2616	EVR---KALKFKACSRRKPDPDALTKTKYLISSYNCPSPPADG---RLQYPGDSGSLH	2668
Oy	588	-----QSBSFTGTGDIINSASALN-----REGILNNAK-----TTSVMIDLPLN	625
Db	2669	SASBSGSGSQSYIPFLIREESTLNPGQVPYPGIGDPSGLEEGQAQHDPPTDSDDLSE	2728
Oy	626	GNHGNSTS 633	
Db	2729	DDOSGSTA 2736	
<hr/>			
RESULT_9			
BA13_HUMAN			
ID	BA13_HUMAN	STANDARD:	PRT: 1522 AA.
AC	060242; 060297;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Bra1n-specific angiogenesis inhibitor 3 precursor.		
GN	BA13 OR KIA0550.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. (LONG ISOFORM).		
RC	TISSUE=Fetal brain;		
RX	MEDLINE=98194217; PubMed=9533023;		
RA	Shiotsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;		
RT	"Cloning and characterization of BA12 and BA13, novel genes homologous		
RT	to brain-specific angiogenesis inhibitor 1 (BA11).";		
RL	Cytogenet. Cell Genet. 79:103-108(1997).		
EN	[2]		
CC	SEQUENCE FROM N.A. (SHORT ISOFORM).		
CC	TISSUE=Brain;		

RX MEDLINE-98290545; PubMed-9628581;
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
 RA Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 CC -1- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION AND
 CC SUPPRESSION OF GLOBLASTOMA.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM. MAY BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
 CC HEART. REDUCED EXPRESSION IS OBSERVED IN SOME GLOBLASTOMA CELL
 CC LINES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 CC -1- SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB005299; BAA25363.1; -;
 DR EMBL: AB011122; BAA25476.1; -;
 DR Genev: HNC:945; BAI3.
 DR MIM: 602684; -;
 DR Interpro: IPR000859; CUB_domain.
 DR Interpro: IPR000832; GPCR_secretin.
 DR Interpro: IPR000203; PKD_cys_rich.
 DR Interpro: IPR000884; TSP1.
 DR Interpro: IPR001879; hormn_receptor.
 DR Pfam: PF00002; 7tm_2; 1.
 DR Pfam: PF00090; tsp_1; 4.
 DR Pfam: PF01825; GPS; 1.
 DR Pfam: PF02793; HRM; 1.
 DR SMART: SM00303; GPS; 1.
 DR SMART: SM00008; HORMR; 1.
 DR SMART: SM00209; TSP1; 4.
 DR PROSITE: PS01180; CUB; 1.
 DR PROSITE: PS0221; GPS; 1.
 DR PROSITE: PS00649; G_PROTEIN_RECP_F2_1; FALSE_NEG.
 DR PROSITE: PS00650; G_PROTEIN_RECP_F2_2; FALSE_NEG.
 DR PROSITE: PS50227; G_PROTEIN_RECP_F2_3; 1.
 DR PROSITE: PS50261; G_PROTEIN_RECP_F2_4; 1.
 DR PROSITE: PS50092; TSP1; 4.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
 KW Repeat; Alternative splicing.
 FT SIGNAL 1 24
 FT CHAIN 25 1522
 FT DOMAIN 25 880
 FT TRANSMEM 881 901
 FT DOMAIN 902 910
 FT TRANSMEM 911 931
 FT DOMAIN 932 939
 FT TRANSMEM 940 960
 FT DOMAIN 961 981
 FT TRANSMEM 982 1002
 FT DOMAIN 1003 1023
 FT TRANSMEM 1024 1044
 FT DOMAIN 1045 1098
 FT TRANSMEM 1099 1119
 FT DOMAIN 1120 1125
 FT TRANSMEM 1126 1146
 FT DOMAIN 1147 1522
 FT TRANSMEM 1523 159
 FT DOMAIN 159 344
 FT TRANSMEM 345 399
 FT DOMAIN 399 399

FT DOMAIN 400 454 TSP TYPE-1 3.
 FT DOMAIN 455 509 TSP TYPE-1 4.
 FT DOMAIN 516 868 GPS.
 FT DOMAIN 942 945 POLY-THR.
 FT DOMAIN 1173 1176 POLY-SER.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 828 828 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 937 937 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 643 665 MISSING (IN SHORT ISOFORM).
 FT VARSPLIC 990 1007 LPAIVATSVGFRTTGY -> KHIDIPHALKMTLNTH
 FT VARSPLIC 1007 1007 (IN SHORT ISOFORM).
 FT VARSPLIC 1008 1522 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 1522 AA; 171490 MM; D22D0A5D4BB62502 CRC64;
 Query Match 12.7%; Score 578.5; DB 1; Length 1522;
 Best Local Similarity 22.7%; Pred. No. 1,3e-29;
 Matches 225; Conservative 163; Mismatches 342; Indels 263; Gaps 39;
 QY 5 RNHLNAGDITTVSRAMQVGLDVOVRNLTTPGSKDA---ARSLNKAWEYVNNLLQP 61
 DB 608 RKNFYAGDLMASVEI-----LRNVTDFKRSYIPASDGVNFFOYISNLDEE 656
 QY 62 ALNARDLTTSQDLRAATM-LIHTVESAVVLDNLKTDIVRENTDNIKLEVARLSTEG 120
 DB 657 NKEKWE---AQQLPSIELMQVIEDFHHVGMGMDFQNSYMTGCVNVAISQKLPAA 713
 QY 121 NLEDKFPENMGHGSTIQLSANTLKKONGEIRVAE----- 157
 DB 714 VLTIDINPMKGRK-----NVDWARSSEDEVIPKSIETTPVSSKEIDESSVEYIGA 764
 QY 158 VLYNNLGPYLTENASMKLGTLEALSTNHSIVNSPVTAIINKFSN-----KYLLDPV 212
 DB 765 VLYKNDLILPT-----LRNYIVNSKIIVTIRPEPKTDSLELELA--- 808
 QY 213 VETVKKIKOSEENFNPCSEFWSYKRTMT-GYWSGCRLLTYKTKTTSCNHLTFEAV 271
 DB 809 -----HL--ANGTLNIPYCVLMDSKTNESLGTMSTGCKVYLDASHTKICLDRLSFPAI 861
 QY 272 LMAHVEYKHSDAVHDLDD-----VITWGIILSLVCL-ICITFCFRGLQSDRNTI 324
 DB 862 L-----AQOPREIIMESGTPSVTLIVSGSLCALITLAVVVAALRYISERSII 913
 QY 325 HKNLCISLFAVELFLIGINRTDQPIACAVEFAALHFFPLAFAFWMLLEGQVLYIMVEV 384
 DB 914 LINCILSISSNILLIVGQYOTANKSICITTTAFLHFFLAFSCVLTLEAMQSMATVTK 973
 QY 385 FESEHSRRKTYFLVCGMPALIVAVSAA-VDRYSYGTDKVCWRLDYTFITSEFGPATLI 443
 DB 974 IRTLILIKRR-FLICGWSGLPALVAVTSVGFRTKGYGDHRCWLSLEGGLYAFVGAAPAAV 1032
 QY 444 IMLNVITGLTALY-----KMFHHTA--ILKESG-CLDINISWYIGATL---- 485
 DB 1033 VLYVMV-IGILVFNKLVSRDGLDKKLLHARGOSEPHSGILLCAACGAVSTALSATY 1091
 QY 486 -----LCILGLTMAFGWLYINES-VVIMAYLFTFINSLOGMFIFEHCVLO 530
 DB 1092 ASNMASLWSSCVVLPALALWMSAVLAMDKRSILQLIFAVFDSLQGEVYVWVHCLR 1151
 QY 531 KKYRREYKCLRTHCSCGKSTESSIGSGKTSGRTPRGVSTGSOSRIIRMMNDTVRQSE 590
 DB 1152 REVQDAF-RCRLRNC-----GQPINADSS 1174
 QY 591 SSFTTG-----DINSSASLNRE-GLLNARDTSVMDTLPLNGN-----HG 629

```

Db 1175 SSFPMHQAQIMTDFEKDVIACRSVLAHKDIGPCRAATITGLTSLISLNDDEEKGTPNG 1234
Oy 630 NSYSIAGCYLSNCVOIIDRGVYHNHNETALEKIKELTNYIPSYLNN-----HERSSE 683
Db 1235 LSTSLTPGVNIS-----KVIIOPTGLHMSMNSNELSPCLKKENSE 1276
Oy 684 QNRNLMNKLVLNMLGSG-----RDDAIVLDDATSFNN-----EES--LGLLILH 725
Db 1277 LRRTYVLCITDDNLGADMDIVHPQERMESDYIVMPRSSVNNQPSMKESKMGIMETLP 1336
Oy 726 EESDAPLLPPRYSTENHP-----HHYTRRRIQDHSSE--FPPLLTNHEEDLOSP 776
Db 1337 HER---LTHYKVPENMNPVWQDFNMNLEQHLAPQEHMONLPFEP-----RTA 1383
Oy 777 HRDSLTYSPTLAGVATSVTTSQTETPPAKCGDAEDVYKSMNGLSRNHVHOLH-- 834
Db 1384 VKNFMASLEDDNAGLSRSEGSTISWSLERRKSRYSDLDFEKVM-----HTRRHME 1436
Oy 835 TYVQLGR--GSSDGFIVPPNKDGTPEGSSKGF 865
Db 1437 LFQELNQKFOFLDRFRDIPNTSSMENPAPNKNP 1469

```

RESULT 10

```

ID CLRS_RAT STANDARD; PRT; 3113 AA.
AC 088278;
DT 15-JUN-2002 (rel. 41, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Cadherin EGF LAG seven-pass G-type receptor 3 precursor (Multiple
GN epidermal growth factor-like domains 2).
OS CELSR3 OR MEGR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98360089; PubMed=9693030;
RA Nakajima M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT Identification of high-molecular-weight proteins with multiple
RT EGF-like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
CC -|- FUNCTION: Receptor that may have an important role in cell/cell
CC signaling during nervous system formation.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- TISSUE SPECIFICITY: Expressed in the brain. Expressed in
CC cerebellum, olfactory bulb, cerebral cortex, hippocampus and
CC brain stem.
CC -|- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -|- SIMILARITY: CONTAINS 9 CADHERIN DOMAINS.
CC -|- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -|- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.
CC -|- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAIN.
CC -|- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB011528; BAA32459.1; -.
CC HSSP; P00740; 1EDM.
DR InterPro: IPR000152; Asx hydroxyl.
DR InterPro: IPR002126; Cadherin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.

```

```

DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000832; GPCR_secretin.
DR InterPro: IPR001879; hormn_receptor.
DR InterPro: IPR002049; laminin_EGF.
DR InterPro: IPR001791; laminin_G.
DR InterPro: IPR000203; PKD_cys_rich.
DR Pfam: PF00002; 7tm_2; 1.
DR Pfam: PF00028; cadherin; 9.
DR Pfam: PF00008; EGF; 6.
DR Pfam: PF01825; GPS; 1.
DR Pfam: PF02793; HRG; 1.
DR Pfam: PF00054; laminin_G; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00180; EGF_Lam; 1.
DR SMART; SM00301; EGF_like; 6.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00282; Lamg; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00232; CADHERIN_1; 7.
DR PROSITE; PS50268; CADHERIN_2; 8.
DR PROSITE; PS00022; EGF_1; 6.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS50221; GPS; 1.
DR PROSITE; PS50025; LAM_G_DOMAIN; 2.
DR PROSITE; PS00649; G_PROTEIN_RECP_F2_1; FALSE_NEG.
DR PROSITE; PS00650; G_PROTEIN_RECP_F2_2; FALSE_NEG.
DR PROSITE; PS50261; G_PROTEIN_RECP_F2_3; 1.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein;
KW EGF-like domain; Calcium-binding; laminin EGF-like domain; Repeat;
KW Developmental protein; Hydroxylation; Signal.
FT SIGNAL 1 31
FT CHAIN 32 313
FT FT 32 2538
FT FT 2539 2559
FT FT 2560 2570
FT FT 2571 2591
FT FT 2592 2599
FT FT 2600 2620
FT FT 2621 2641
FT FT 2642 2662
FT FT 2663 2679
FT FT 2680 2700
FT FT 2701 2724
FT FT 2725 2745
FT FT 2746 2752
FT FT 2753 2773
FT FT 2774 3313
FT FT 317 424
FT FT 425 536
FT FT 537 642
FT FT 643 747
FT FT 748 849
FT FT 850 952
FT FT 953 1058
FT FT 1059 1160
FT FT 1160 1257
FT FT 1257 1366
FT FT 1366 1424
FT FT 1426 1462
FT FT 1466 1505
FT FT 1506 1710
FT FT 1710 1749
FT FT 1749 1935
FT FT 1937 1972
FT FT 1973 2011
FT FT 2012 2044
FT FT 2046 2081
FT FT 2087 2120
FT FT 2120 2527
FT DOMAIN 2475

```

Query Match	Best Local	Similarity	Score	DB	Length
Matches 194: Conservative	132: Mismatches	334: Indels	174: Gaps		
47 NKAMEYVNNLLQPOLANWRDL	---TTSQDLRAATMLHTFEESAFVLADNLKTDIVR	103			
2254 NENLWAGSALLAPETGDMALGCRAPSGSGSAGLVNH	-LEETVAATLARNN--DLTY	2309			
104 EN----	TDNIKLEVARL---STEGCNLEDLKFPENMGH-----	134			
2310 LNPVGLVPNNIMLSDIRMBOPSSGQAHRYPRYNSLFRGQDAMPDRHTVLLPSQSPQS		2369			
135 ---STIGLSANT-----	LKQNGRNGEIRFAVLVYNNIGSPYUJSTENASKLCT	178			
2370 PSEVLPRTSSNANATASGVSPAPRLEPSEPGISIVILLVYRALGGLPAFOAERRRA		2429			
179 EALSTNHASIVYSPITAIANKFEFSKVVLAADVFTVK	---HIKQSEENFNPNCSFWS-	234			
2430 R---LPQNVNMSPVVSAV--	FRGRNLRGALVSPNLEFRLLQJATNRSKAICQWMP	2483			
235 YSKRTMTGWSYTOGRLLLTNTKHTTSCSNHLTNFAVLMAHVEVKSIDAVNLD	-LIDVYT	293			
2484 PERADQHGMMTRADELVIRNRSNHARCRCRSTGTCGVLMADAPREBLEG	-DLELLAVTT	2541			
294 WVGILLSLVCLLICITFTFCFGRGLOSDRNTIKNLCLISLFEALLFLIGINPTDPIACA		353			
2542 HVVVAASVATVILTAVALVLLSLKSNVNGIHANVAAALGVALLFLGIGIRHTNQDLCT		2601			
354 VFAALLHFFFLAFLFMMFLFEGVQILTMVLEVEESESRRKYYLVVYGCPALIVANSAV		413			
2602 VAILLHFFEFSTFMMFLVQGLVLRMOVQEPNNVNGARFYNHAGVGVPAVLGLAVGL		2661			

OY	414	DYRSAGTDKVCWMLDQVTFMISFGRPTLIMLVFLGALYKPMHNAIKLPSSGD	473
OY	414	: : :	
Db	2662	DPEGNDPFCWISIHBPILMSFAGPIVLIVNNGI-----MFLAARTSCSGORE	2713
OY	474	NIKSNVI----GATALLCLGLTWAGLMYNESFTVMAYLFETFNLSLOGMFIFHCYL	529
Db	2714	AKTISTVLRTHSSPLLLLVASAMLGSLAVNHNSVLAIPHNLHAGLCGGVLVLLFCYL	2773
OY	530	OKKYRKEFGKCLRPHCCSGSTESSIGSGTSSSRIPGRYSTGSOSRIIRMMNDTVRKOS	589
Db	2774	NADARAAM-----TPALGKKAPE-----ETRAPPBG-----	2802
OY	590	ESSFTIGINSASINREGLLNNARDPDYVDLT--PLNGHNHSYSIASGEYLSMCYOII	647
Db	2803	: : : : : : : : : : : : :	
OY	648	DRGYNHMETALEKKILKELTSNYIPSYLNHHSSBQNRLNKLVNNGSREDREDAIVL	707
Db	2857	---RHGSTA-----EHAESHLOAHAGPTOLDVAMFHR-----DACADS---	2893
OY	708	DDATSFNHEBSLGLIELHESDA-----PLLPRPVSTEMHQPHNTRRTRIQDISE	759
Db	2894	DSDSLSEEBERSLIPSSSEDNQRTGRFORPL--BRAQSEELLAAH-----PKDVDG	2944
OY	760	---SFPLL-----TNEHTEDLOSHPHDLSYTSMPTLAGVAATES	796
Db	2945	NDLSTYPALGECBAACPALQANGSERRLGDS-NKDAANNOPELALTSGDET	2997
<hr/>			
RESULT 11			
CLR3_MOUSE	ID	CLR3_MOUSE	STANDARD; PRT; 3301 AA.
AC	091ZIO: 09ESD0:		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Cadherin EGF LAG seven-pass G-type receptor 3 precursor.		
GN	CELSR3.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGES.		
RC	STRAIN=C57BL/6;		
RX	PubMed=11850187;		
RA	Tissir F., De-Becker O., Goffinet A.M., Lambert de Rouvroit C.A.;		
RT	"The flamingo-related mouse Celser family (Celser1-3) genes exhibit		
RL	distinct patterns of expression during embryonic development.";		
RL	Mech. Dev. 109:91-94(2001).		
RN	[3]		
RP	TISSUE SPECIFICITY.		
RX	PubMed=10790539;		
RA	Chromstone C.J., Barclay J., Rees M., Little P.F.R.;		
RT	"Chromosomal localization of Celser2 and Celser3 in the mouse: Celser3 is		
RL	a candidate for the floppy (fip) lethal mutant on chromosome 9.";		
RL	Mamm. Genome 11:392-394(2000).		
CC	-1- FUNCTION: Receptor that may have an important role in cell/cell		
CC	signaling during nervous system formation.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-1- TISSUE SPECIFICITY: Expressed in the CNS and in the eye.		
CC	-1- DEVELOPMENTAL STAGE: Predominantly expressed in the CNS, the		
CC	emerging dorsal root ganglia and cranial ganglia. In the CNS,		
CC	expression is uniform along the rostrocaudal axis. No expression		
CC	is detected until somite stages. Between E10 and E12, expression		

ID	CLRS3_MOUSE	STANDARD;	PRT;	3301 AA.
AC	Q91Z10; Q9ESD0;			
AD	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Cadherin EGF LAG seven-pass G-type receptor 3 precursor.			
GN	CELSR3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RY	[1]			
RP	SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGES.			
RC	STRAIN=C57BL/6;			
RX	PUMed=11850187;			
RA	Tissir F., De-Backer O., Goffinet A.M., Lambert de Rouvroit C.A.;			
RT	"Developmental expression profiles of Celser (Flamingo) genes in the			
RL	mouse.";			
RL	Mech. Dev. 112:157-160(2002).			
RN	[2]			
RP	SEQUENCE OF 2574-3046 FROM N.A., AND DEVELOPMENTAL STAGE.			
RX	PUMed=11677057;			
RA	Forrestone C.J., Little P.F.R.;			
RT	"The flamingo-related mouse Celser family (Celsr1-3) genes exhibit			
RL	distinct patterns of expression during embryonic development.";			
RL	Mech. Dev. 109:91-94(2001).			
RN	[3]			
RP	TISSUE SPECIFICITY.			
RX	PUMed=10790539;			
RA	Forrestone C.J., Barclay J., Rees M., Little P.F.R.;			
RT	"Chromosomal localization of Celsr2 and Celsr3 in the mouse; Celsr3 is			
RL	a candidate for the floppy (fip) lethal mutant on chromosome 9.";			
RL	Mamm. Genome 11:392-394(2000).			
CC	-1- FUNCTION: Receptor that may have an important role in cell/cell			
CC	signaling during nervous system formation.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- TISSUE SPECIFICITY: Expressed in the CNS and in the eye.			
CC	-1- DEVELOPMENTAL STAGE: Predominantly expressed in the CNS, the			
CC	emerging dorsal root ganglia and cranial ganglia. In the CNS,			
CC	expression is uniform along the rostrocaudal axis. No expression			
CC	is detected until somite stages. Between E10 and E12, expression			


```

Db 2247 NENLLMAGSALLALETGHLMALQORAPGSGAGLVOH-LBEGYAATLANNMELTYLNP 2305
OY 102 VRENTDNKLEVARL-----STEGNLEDKPEPNHG-----S 135
Db 2306 VGLVTPNIMLSIDMEHBSSTOGARRPRYSNLFPGODADPHTHVLLPSQASOPSE 2365
OY 136 TIOLASANT-----LKONGRCEIRAFVLYNNIGPYLSTENASMKLGTAL 181
Db 2366 VLPRTSNANENATASSVSPAPLPESEPGISIVLLYRALGQIPQOAFERRGAR- 2423
OY 182 STNSVIVNSVYTAALINKESKNVYLAADPVYTVK---HIKQSEENPNPCSWYSKR 238
Db 2424 -LPONPWNSPVVAV--FHGRNPLGVLVSPINEEFLQLQANRSKALCYOMDPGP 2479
OY 239 T-MGYSTOGCRLLTTHKTHHTSCNHLTFEALMAHEVXKSDAYDL-LDVIYTVG 296
Db 2480 TDQGMGTARCELVHRNGSIARCRSTGTGYLMDASPERLEG--DELLAVFTHV 2537
OY 297 ILLSLVCLLICIFTECFEFGIQSDRNTIHKMLCISLFAVELLFTGINRTDQIACAVFA 356
Db 2538 VAVSVTALVTAVALLSIRSLKSNVRGIHANVAALGVAELLFLGIRHTNQLCTAVA 2597
OY 357 ALLHFFPLAATMFLBEVOIYIMLVEFESEHSRRKFTYVCGMPALIYAASANDYR 416
Db 2598 ILLHYFLSTFAMLLVGLHLYRMQOVERNDRGAMREYHAIQGVPAVLLGLAVGLDPE 2657
OY 417 SYGDKCWLRLDYTFISFGPATLILMLNVIFGLALYMFHNTALPKESGLDNK 476
Db 2658 GYGNPDCMWSIHPLRMSFPGIYLVYVNN-----GIMFLAARTSCSTGOREAKK 2709
OY 477 SWVI--GAIALCLGLTMAFGMYINESTVIMAYLETIFNSLOGMEFIFFHCYLOKKV 533
Db 2710 TSVLTIRSSFLILLVLSASWIFGLAVNHSILAFHYLHAGCGIQLGLAVLLFCVLANDA 2769
OY 534 RKEGKCLRTCCSGKSTRESSIGSGKTSGRTPGRYTSQSQRIRRMANDYVRKQSESEF 593
Db 2770 RAAW-----TPACLGKKAPE---ETRPAPGP----- 2794
OY 594 ITGDISSASLNREGILNARDTSVMDL--PLNGHNSYSIASGEYLSNCOIIDRGY 651
Db 2795 -SGAYNNATLAFEEBSGLIRITIGASTVSSVSARSGRADODOSQGRSLIRBNLV- 2848
OY 652 NHNETALEKKILKELTNSYIPSYLNHNRSESONRNMLKNLNLGSGREDDATVLDAT 711
Db 2849 RHGSTA-----EHTERSLOAHAGPTDLVAMFHR-----DAGADSDS--DSDL 2889
OY 712 SFNHEESIGLELHESDA-----PLLEPRVYSTENHOPHHYTRRRITQDHSE--- 759
Db 2890 SLEBERSLSTPSSSEEDNGRTGRFRPL--RAAQSERLLAH-----EKDVGNDLL 2940
OY 760 SFPL-----TNEHTEDLOSPPRDSLYTSMPTLAGVAATES 796
Db 2941 SYWPALECEAPCALQAMGSERLGLDS-NKDAANNQPELALTSGDET 2989

```

```

RX MEDLINE-202025599; PubMed-10716726;
RA Wu Q., Maniatis T.;
RT "Large exons encoding multiple ectodomains are a characteristic
RT feature of protocadherin genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:3124-3129(2000).
RN [2]
RP SEQUENCE OF 1954-3312 FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-98360089; PubMed-9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple
RT EGF-like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
CC -1- FUNCTION: Receptor that may have an important role in cell/cell
CC signaling during nervous system formation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 9 CADHERIN DOMAINS.
CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF231023; AAF61929.1; -
DR EMBL; AB011536; BAA32464.1; -
DR HSPB; P00740; IEDM.
DR Genew; HGNC:3230; CELSR3.
DR MIM; 604264; -
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR000832; GPCR_secretin.
DR InterPro; IPR001879; hormn_receptor.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR000203; PKD_cys_rich.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF00028; cadherin; 9.
DR Pfam; PF00008; EGF; 6.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00054; laminin_G; 2.
DR PRINTS; PR00249; GPCRSECRETIN.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00112; CA; 8.
DR SMART; SM00180; EGF_lam; 1.
DR SMART; SM00001; EGF-like; 6.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00282; Lamg; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00232; CADHERIN_1; 7.
DR PROSITE; PS0268; CADHERIN_2; 8.
DR PROSITE; PS00022; EGF_1; 6.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS00221; GPS; 1.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR PROSITE; PS0227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS0261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS00025; LAM_G_DOMAIN; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 1.

```

```

RESULT 12
CLR3_HUMAN
ID CLR3_HUMAN STANDARD; PRT; 3312 AA.
AC Q9NYQ7; 075092;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Cadherin EGF lag seven-pass G-type receptor 3 precursor (Flamingo
DE homolog 1) (hfm1) (multiple epidermal growth factor-like domains 2)
DE (epidermal growth factor-like 1).
GN CELSR3 OR COFILIN OR FMIL OR EGFL1 OR MEGF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.

```

FT	CARBOHYD	1649	1649	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1713	1713	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	1770	1770	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	2053	2053	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	2177	2177	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	2196	2196	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	2386	2386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	2474	2474	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	2506	2506	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CONFLICT	2158	2158	G -> GLRGAG (IN REF. 2).	(POTENTIAL).
SO	SEQUENCE	3312 AA;	358251 MW;	BE208703651AA45 CRC64;	
Query Match 11.1%; Score 507.5; DB 1; Length 3312;					
Best Local Similarity 22.3%; Pred. No. 1,7e-24;					
Matches 237; Conservative 158; Mismatches 422; Indels 245; Caps 38;					
OY	6	NHLNAGDITTSVRAMDOLVGLLDVQRLNLTGEGKDSAAKSLINKAMVETVNNLLOPOLNA	65		
Db	2220	DHYSDQDVRVTAARLNLHLAEESHQOQGLTATODA--HPENMLMGASLLAPETGDL	2276		
OY	66	WRDL---TTSQDLRAATMLTHTVEASAVLADNLKTDI--VRENTDRIKEVARL----	116		
Db	2277	WALAQARPGSGSPSAGLVRH--LEBYAATLARNMELTYLANMGVLTPTIMLSIDRMEPS	2335		
OY	117	STEGNLEDLKPEPNNGH-----STIOLASANTLKONGRNGEIR-	154		
Db	2336	SPRGARRPRYHSNLFRRQDAMDPTHVLLRPSQSPRPSVPLTPSSIEKSTSSVYPP	2395		
OY	155	-----VAFVLYNNIGPYLSTENASMKIGTELSTNHSYVNSPVITAI--NK	200		
Db	2396	PAPERPEGISIIILLVYRTLGGLLPQFOAERRGAR---LPONVYNSPVYSAVFGHR	2452		
OY	201	EFSNKVYIADAVVFVKIKIKOSEENFNNGCFWS-YSKRTMGWSTQGRLLTNTKRT	259		
Db	2453	NFLKRI-LESPISLEFR-LIOTANRSKAIQYOMDPGICAEQGWTAARDCELVHRNGSHA	2510		
OY	260	TCSCHLNTNFAVLAHAEVFKSHDAVHD-LDVITWVIGLLSLVCLICITTFEFGRIQ	318		
Db	2511	RRCGRGTGCVGLMDASPRELEG--DELLAVFTHVVANVAALVYLAAILLSLRSLK	2568		
OY	319	SDRNTIKNLCISLFEVALLFLIGINTDQPIACAVFALLHFFFLAFTMPLLEGVOLY	378		
Db	2569	SNVRGIIHANNVAALGVALLFLGIIHRTNOLVCTAVAILLHYFELSTFAMLFVQGLHLY	2628		
OY	379	IMLVFPESEHSRRKYEFLVYGMPALIVANAVANDVSYCTDKVCWMLRTTYFLWSIG	438		
Db	2629	RMOVEPRANDVGRAMFEYHALDMGVAVLLGLAVGIDPEGVNPDPCWISVEHPLWASG	2668		
OY	439	PATLIMLVNVLFLGIALTKMFHHTAILEPESGCDLNISWYI---GATALLCLGLTWAF	495		
Db	2689	PVYLVIYVN-----GTMFLLAARTSCSTGQRAKKTSALTLSRSPLLLLLYSASLFL	2740		
OY	496	GLMVTINESTVIMAYLFTIFNSLOGMFIPIECVLOKKVKEY-GKCLRTHCS-----	547		
Db	2741	GLLAVNHSILFAHYLHAGLGLGGLAVULLFLCVLADARAAMPACLCGRKKAPEARPA	2800		
OY	548	-----GKSTSSIGSGKTSSTPRPKRYSTGSGRIRRMNDT	584		
Db	2801	GLGPGAVNNALFEESGLIRITLGLASTVSVSARS--GRQDDOSQGRNSYLRD--NVL	2856		
OY	585	VKROSESSF-----LTGDISSASINREGILN-----	611		
Db	2857	VRHGSAAHDTSHLOAHNGPDLDYAMMHRDAGADSDSDLSLEERSLSIPSESDN	2916		
OY	612	-----NARDISVMDLP--LNGHNGNSYIASGE-YLSNC-----V	644		
Db	2917	GRTGRFORPLCRAAOSERLLTHPRKDVGNLDSWPALGCEAPCALQWGSERRLGL	2976		
OY	645	QIIDRGYVHN-----ETAL-----EKKILK-ELTISNYIPSYNNHRSSEQRN	687		
Db	2977	DTSKDAANNNDPALRTSGDETSLGRACORRKGILKNLQYLPVQYTRGAEELSCRAAT	3036		
OY	688	LMNKLVNNIGSGGR-----EDDAIVLDDATSPFHNEESIGLELH-----EESDAPLLP-	735		

```

Db 3037 LGHRAVPAASGRIVAGGTGTSQOPASRYSSREQLDLLRLRQLRSREPLEAPAPVLRPL 3096
QY 736 -RYVSTP-----NQHPIHYTRRIPODHSEFFPLTNHEHEDQSPHSDLYT 783
Db 3097 SRGSGOCMDAPGRLEPPKRGSTLPRKPPDPGAMAGRFGSRDALDAP-REWLST 3155
QY 784 SMPTLAGVAATESVTSTGTCTEPPAKGCD---AEDVYKSM-----NLGSRNHVHQL 833
Db 3156 LPPRR-----TRLDQPPPLPLSPQLSRDLPLSRPLDSLRSRSNSRQDLQV 3207
QY 834 HTYQQLGRSSSDGFIVPNKDGTPP-----EGSSKGPAN 867
Db 3208 PSRH-----PSREALGPLQLRLARSDSVSGPSH 3236

RESULT 13
BA12_HUMAN STANDARD; PRT; 1572 AA.
ID BA12_HUMAN
AC 060241;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brain-specific angiogenesis inhibitor 2 precursor.
GN BA12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA MEDLINE:98194217; Pubmed-9533023;
RX Shitouchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
RT "Cloning and characterization of BA12 and BA13, novel genes homologous
RL Cytochrome C gene. 79:103-108(1997).";
CC - FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
CC HEART, THYMUS, SKLETERAL MUSCLE, AND DIFFERENT CELL LINES.
CC - SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC - SIMILARITY: CONTAINS 4 TSP TYPE-1 DOMAINS.
CC - SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB005298; BAA25362.1;
CC Genew: HGNC:944; BA12.
CC MIM: 602683;
CC InterPro: IPR000832; GPCR_secretin.
CC InterPro: IPR000203; PKD_cys_rich.
CC InterPro: IPR000884; TSP1.
CC InterPro: IPR001879; horma_receptor.
CC Pfam: PF00002; 7tm_2; 4.
CC Pfam: PF00090; tsp_1; 4.
CC Pfam: PF01825; GPs; 1.
CC Pfam: PF02793; HRM; 1.
CC SMART: SM00303; GPs; 1.
CC SMART: SM00008; Hormr; 1.
CC SMART: SM00209; TSP1; 4.
CC PROSITE: PS50221; GPs; 1.
CC PROSITE: PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
CC PROSITE: PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
CC PROSITE: PS50227; G_PROTEIN_RECEP_F2_3; 1.
CC PROSITE: PS50261; G_PROTEIN_RECEP_F2_4; 1.
CC PROSITE: PS50092; TSP1; 4.

```

```

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
Repeat.
FT SIGNAL. 1 20
FT CHAIN. 21 1572
FT DOMAIN. 21 924
FT TRANSMEM. 925 945
FT TRANSMEM. 946 953
FT TRANSMEM. 954 974
FT TRANSMEM. 975 982
FT TRANSMEM. 983 1003
FT TRANSMEM. 1004 1024
FT TRANSMEM. 1025 1045
FT TRANSMEM. 1046 1066
FT TRANSMEM. 1067 1087
FT TRANSMEM. 1088 1141
FT TRANSMEM. 1142 1162
FT TRANSMEM. 1163 1189
FT TRANSMEM. 1189 1190
FT TRANSMEM. 1190 1572
FT TRANSMEM. 1572 351
FT TRANSMEM. 352 406
FT TRANSMEM. 407 461
FT TRANSMEM. 463 517
FT TRANSMEM. 517 911
FT TRANSMEM. 911 122
FT TRANSMEM. 122 177
FT TRANSMEM. 177 225
FT TRANSMEM. 222 1303
FT TRANSMEM. 1303 1358
FT TRANSMEM. 1352 1418
FT TRANSMEM. 1413 1418
FT TRANSMEM. 94 94
FT TRANSMEM. 179 179
FT TRANSMEM. 180 180
FT TRANSMEM. 344 344
FT TRANSMEM. 425 425
FT TRANSMEM. 548 548
FT TRANSMEM. 633 633
FT TRANSMEM. 855 855
FT TRANSMEM. 1572 AA; 171140 MW; A975645B77BC285 CnC64;
SO SEQUENCE

Query Match 11.18; Score 505.5; DB 1; Length 1572;
Best Local Similarity 21.88; Pred. No. 7.7e-25;
Matches 233; Conservative 146; Mismatches 347; Indels 345; Gaps 44;

QY 10 AGDITVSRAAMDVLGLVDVLRNLT-----PGKDSAAANSLKAWETVNNLQIP 60
Db 621 SGDLFFSV-----DI-LRNVDTFKRATVYPADV-----QRFQVVSFWDA 663
QY 61 QALNAMDLTTSQDLRAATM-LHTVE-----ESAFVLADNLT----- 97
Db 664 ENKEKMD--AQGVSPGVHLRVEDEFLHVGDAKARQSSLIYVDNLVISIOPRVS 720
QY 98 -----KDIYRENTDNKL--EVARLSTEGNLEL-----KF 127
Db 721 AVSSDITFPNMRGRGMKQWVRHSHEDRLFLEPREVLSLSPGKPAISGAAGSPGRGPGTV 780
QY 128 PENNGHSTIOLSAANTLKQNGRNGEIRVAFVLYNNLPYSTENASMKLGTEALSTNHSV 187
Db 781 PRGCHSHQRLRPADP-----DESSYFVIGAVLYTTLGLLPPRPPL----- 823
QY 188 IVNSPVITAAINKEFSKNKYIADPVFTVKHIKQSEENFNPNCSFMSYSKR-TWTGWST 246
Db 824 AVTSRVTVVVRPPTPP--AEPLI-TVELSYIINGTTPHCASMPYSRADSSSGMDPT 879
QY 247 QGCRLLTNTKHTTSCNHLTNFAVLAHVEYKKSADAVHDLIDVITWVGLLSVCLLI 306
Db 880 ENCQTLFQAHTRCQCOHSTFAVL-----AQPKDLTLELASSPVLYIGCAVS 931
QY 307 CTFTF-----CFRGLSDRNTIHNKLSLFEVAFLEFLIGINTDQPIACAVFAAL 359
Db 932 CMALLTILAIYAAFWRIKSRSTIILNFCILSIANSILILVQGSRYLSKGVCTMTAFL 991
QY 360 HFFFLAFTWMTLEGVOLYIMLVEFESEHSRRRYFVLGVGMDALIVAVSAA-VDYRSY 418

```



```

FT SITE 231 233 CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN 1365 1584 NECESSARY FOR INTERACTION WITH BAP1.
FT 1584 1584 INDISPENSABLE FOR INTERACTION WITH BAP1.
FT CAROHD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 844 844 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 877 877 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CAROHD 881 881 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1584 AA: 173531 MM: DEAB8F28C7874513 CRC64;

Query Match 10.9%; Score 496.5; DB 1; Length 1584;
Best Local Similarity 25.5%; Pred. No. 36-24;
Matches 175; Conservative 123; Mismatches 268; Indels 121; Gaps 32;

OY 3 QRRNHADDTYSVRAMQVGLLDVQLRNL-----TPGKDSAAASLKAW 51
DB 667 QRLVEISQDGTSTST---GDLLSTIDV-LRNMEIFRAYSPEDDVON-----FY 713
OY 52 ETVNNLLOPQALNARDLTSTDLRA--ATMLHTVEESAFVADNLKTDIYRENTDI 109
DB 714 QLSNLAENRDKMEA---QLAGPNKELFRLVDEFDVIGFRMKDLRDAYQYTDNL 769
OY 110 KLEVARLSTEGNLEDLKPEPNMGHS-----TQLS--ANTLQNGRNGEIRVA 156
DB 770 VLSIHLKPLASG-ATDISFPMKGWRATGDAKYPEDRVTYSKVSFGLTEADBSAFVVG 828
OY 157 FVLYNNLGPYLTSTENASMKLTGATLSTNSVYNSVITRAIKESNKKYVLADPVVTV 216
DB 829 TVLYRNLGSLAQ-----RNTTVNSKVISTVAPPPRS---LRTPLEIEF 872
OY 217 KRIKQSEENFPNCSFW-----SYSKRTWTGYSTQGCRLTNTKTHSTCNHLTNEFA 270
DB 873 AHMYNG--TTCNCILMDENDVPSSAPQLGPMWNRGCTVLDLARTCLDRSTFA 930
OY 271 VLM-----AHVEVKSDAVHDLIDVTVWGIILSVCLLICITFCFGRGLOSDRNTIH 325
DB 931 IIAQLSADANMERATLPSTV---TLIVGCV-SSTLLMLVITVYVSVYIRSEERSVIL 984
OY 326 KNCISLFAVELLEFLIGINTDPIACAVPALHFEFLAFLMFLLEGVQLIMLEVFE 385
DB 985 INCLSIITSSNALITIGOTFRNKVKCTLVAALHFEFLSSFCWVLTETAMQSTMAV---- 1040
OY 386 ESEHSR-----RKTYFLVGYGMPALIAVNSA--DYRSYGTDKCMLRLDTYMSFIPA 440
DB 1041 -TGHNLRLIRKRFELGMLPALVAISGFCFKAGYSMTMNCWLSLEGILYAFVPGA 1099
OY 441 TLIIMLVNIFLGIALY-KMFHHTAT-----LKPSGCLDNIKSVIGAILLCLGLTWAR 495
DB 1100 AAVVLYNMV-IGILVFNKLVSKDGIIDKILKERAGA---SLW--SSCVVLPPLATLWMS 1152
OY 496 GLMYINE-STVIMAYLFTTFNSLOGMFIFFHVLCKYKVEKCKLCRTCCSGKSTRESS 554
DB 1153 AVLAATDRNSALFQILFAVFDSEGFYVWHLKLRREV-QDVVKCRV-----DROE 1204
OY 555 IGSQSGSGSRTPG--RYSYTSQGRIRRMNDYTRKO---SESSFITGINSASLNREGI 609
DB 1205 EGGDGGSGGQNGHQALMDTFEKDVLDACSYVANKDIAOCTATITGTILKRPPLPEBEKL 1264
OY 610 -LNNARDTSV-MDTPLNGN-----HGN 630
DB 1265 KLAHAKGPTNFNSLPAVNSKHLHGS 1291

RESULT 15
STAN_DROME STANDARD: PRT: 3579 AA.
AC Q9V5N8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin-like wing polarity protein stan precursor (starry night

```

```

DE protein) (Flamingo protein).
GN STAN OR FMI OR CG11895.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RC TISSUE=Embryo;
RX MEDLINE=20025940; PubMed=10556066;
RA Chae J.W., Kim M.-J., Goo J.H., Collier S., Gubb D., Charlton J.,
RA Adler P.N., Park W.J.;
RT "The Drosophila tissue polarity gene starry night encodes a member of
RT the protocadherin family."
RL Development 126:5421-5429(1999).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLINE=99418630; PubMed=10490098;
RA Usui T., Shima Y., Shimada Y., Hirano S., Burgess R.W., Schwarz T.L.,
RA Takeichi M., Demura T.;
RT "Flamingo, a seven-pass transmembrane cadherin, regulates planar cell
RT polarity under the control of frizzled."
RL Cell 98:585-595(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosser C., Gabrielian A.E., Gary N.S., Gilbert M.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: Involved in the fz signaling pathway that controls wing
CC tissue polarity. Also mediates homophilic cell adhesion. May play
CC a role in initiating prehair morphogenesis. May play a critical
CC role in tissue polarity and in formation of normal dendrite
CC fields.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 11:11:24 : Search time 13 Seconds
(without alignments)
1089.484 Million cell updates/sec

Title: US-09-744-226a-1

Perfect score: 4563
Sequence: 1 AEQTRNHLNAGDIRYSVRAM.....KDGTPPECCSKGPAHLVLSL 872

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PC01_NEM_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1088.5	23.9	645	10 US-09-796-338A-2	Sequence 2, Appli
2	1088.5	23.9	690	9 US-09-905-291A-49	Sequence 49, Appli
3	1088.5	23.9	690	10 US-09-828-366-7	Sequence 7, Appli
4	1088.5	23.9	690	10 US-09-909-320-49	Sequence 49, Appli
5	1088.5	23.9	690	10 US-09-909-088A-49	Sequence 49, Appli
6	1088.5	23.9	713	10 US-09-925-300-1407	Sequence 1407, Ap
7	927	20.3	240	10 US-09-978-486-5	Sequence 5, Appli
8	776	17.0	661	10 US-09-764-853-679	Sequence 679, App
9	776	17.0	661	10 US-09-764-898-224	Sequence 224, App
10	774	17.0	652	10 US-09-992-647-1	Sequence 1, Appli
11	715.5	15.7	886	10 US-09-992-647-11	Sequence 11, Appli
12	700.5	15.4	3034	10 US-09-737-149-25	Sequence 25, Appli
13	700.5	15.4	3034	10 US-09-737-149-30	Sequence 30, Appli
14	692	15.2	3014	10 US-09-737-149-2	Sequence 2, Appli
15	670.5	14.7	1447	10 US-09-808-571A-2	Sequence 2, Appli
16	652.5	14.3	344	10 US-09-992-647-12	Sequence 12, Appli
17	636.5	13.9	2923	10 US-09-788-711A-4	Sequence 4, Appli
18	636.5	13.9	2956	10 US-09-788-711A-2	Sequence 2, Appli
19	635.5	13.9	874	10 US-09-963-766-6	Sequence 6, Appli

20	595	13.0	717	10 US-09-925-300-1299	Sequence 1299, Ap
21	581.5	12.7	733	12 US-10-036-328A-4	Sequence 4, Appli
22	581.5	12.7	1210	12 US-10-036-328A-2	Sequence 2, Appli
23	521.5	11.4	1354	10 US-09-808-571A-4	Sequence 4, Appli
24	520	11.4	3313	10 US-09-737-149-29	Sequence 29, Appli
25	517	11.3	231	10 US-09-978-486-4	Sequence 4, Appli
26	513	11.2	331	9 US-09-860-670-118	Sequence 670, App
27	513	11.2	331	10 US-09-764-853-671	Sequence 671, App
28	513	11.2	331	10 US-09-764-898-219	Sequence 219, App
29	508.5	11.1	714	10 US-09-818-264-2	Sequence 2, Appli
30	507	11.1	513	10 US-09-818-264-4	Sequence 4, Appli
31	500.5	11.0	662	12 US-10-036-328A-8	Sequence 8, Appli
32	500.5	11.0	1138	12 US-10-036-328A-6	Sequence 6, Appli
33	497	10.9	235	10 US-09-978-486-6	Sequence 6, Appli
34	496.5	10.9	448	10 US-09-818-264-5	Sequence 5, Appli
35	488	10.7	848	10 US-09-843-164-6	Sequence 6, Appli
36	474	10.4	568	10 US-09-843-164-2	Sequence 2, Appli
37	463	10.1	1349	10 US-09-747-835A-52	Sequence 52, Appli
38	454.5	10.0	986	10 US-09-747-835A-53	Sequence 53, Appli
39	454.5	10.0	1325	10 US-09-747-835A-24	Sequence 24, Appli
40	454.5	10.0	1346	10 US-09-747-835A-20	Sequence 20, Appli
41	454.5	10.0	1346	10 US-09-747-835A-61	Sequence 61, Appli
42	454	9.9	560	10 US-09-843-164-12	Sequence 12, Appli
43	452.5	9.9	807	10 US-09-895-686-6	Sequence 6, Appli
44	451.5	9.9	1371	10 US-09-747-835A-25	Sequence 25, Appli
45	445	9.8	769	10 US-09-843-164-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-796-338A-2
Sequence 2, Application US/09796338A
Patent No. US20020061522A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.
TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
FILE REFERENCE: 10448-020001
CURRENT APPLICATION NUMBER: US/09/796,338A
CURRENT FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/186,059
PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 645
TYPE: PRT
ORGANISM: Homo sapiens
US-09-796-338A-2

Query Match 23.9%; Score 1088.5; DB 10; Length 645;
Best Local Similarity 41.4%; Pred. No. 1.4e-77;

Matches 227; Conservative 111; Mismatches 189; Indels 21; Gaps 11;

QY	8	LNRGDTTVVRAMDOVLGLLDVQLRNLTPGCKDSARSLNKAKVETVNNLLOPQALNMR	67
DB	109	LSPTDITTYIEILAESSSLG--YKNNITSAKDTLSSTLETKVNNFVORDPFVWD	166
QY	68	DLTTSQDLAAATPMLHTTVEESAFVLADNLKTDIVRENDONIKLEVARLSTEGNLEDAKF	127
DB	167	KLSVNRKRLHLKMTVDAATLRISQOKTTEFTPTSTIDALAKVFFDST-NMKHHIP	225
QY	128	PENMGSTIOLSAITLKNGRGELRAVFLVNLNIGPYL-STENASMKLTGEALSTNHS	186
DB	226	HNM-DGDYINIFPKKAAVDSNGVAVALFYKKSIGPLSSSDNFLKPOVNDSEEE	284
QY	187	VIVNSPVITAAINKESN--KYLADPVVFTYKHKILOSEENPNPCSEFYSKRTTGTW	244
DB	285	RIISS-VIVSWS---SNPPTLYELEKIFFTLSHRKVT-DRYRSICAFNNYSPDTNGSW	339
QY	245	SNQGCGLLTNTKTHPTSCNHLTNFAVLMA---HVEVKSDAVHDLDLVITWVIGLLSL	301

Db 340 SSECCELTYSNETHTSCRNHLTHFAILMSSGSPSIGIKDYN-----ILTRITOLGIITSL 394
OY 302 VCLLICIFTCFPRGLSDNRTIHKNCISLFEVALLFLGINFTDQPIACAVFAALLHF 361
Db 395 ICALICITFTWFESEIOSTRITTIHKNCISLFEVALLFLGINFTDQPIACAVFAALLHF 454
OY 362 FFLAFTWMELEGVQYIMLVEVESESRKRYFYLVGYGMPALIVASAVDRSYGTD 421
Db 455 FFLAFAFMOCIEGHLHLYVGVLYNKGFLHKNFYIFGYLSPAVVGFSAALGRRYGT 514
OY 422 KVCMLRDTYFMSFIPATLIIMLVNIFIGIALYKMFHTAILKPSGCLDNKISWVG 481
Db 515 KVCMLSTENNFMSFIPACILILVNLAFGVIIYKFRHTAGLKPEVSCFENIRSCARG 574
OY 482 AIALCLLGTWAGLWYNVESTYIMAYLFTFNSLQGMFIFPHCVLOKKVREYKCL 541
Db 575 ALALLFLGTWTWIFGVLYHVHVASVYVYALFTVSNAGFMFLFLCVLSKRIODEYTRLF 634
OY 542 R-THCCSG 548
Db 635 KNPCCCG 642

RESULT 2
US-09-905-291A-49
Sequence 49, Application US/09905291A
Patent No. US20020160374A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Malher, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,291A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 49
LENGTH: 690
TYPE: PRT
ORGANISM: Homo sapiens
US-09-905-291A-49

Query Match 23.9%; Score 1088.5; DB 9; Length 690;

Best Local Similarity 41.4%; Pred. No. 1.5e-77; Matches 227; Conservative 111; Mismatches 189; Indels 21; Gaps 11;

OY 8 LMAQDITYSVRAMDQVLGLDVLQRLNLTGPKGDSARSRLKAWETVNNLLQDALNAWR 67
Db 154 LSPDITTYIELLAESSLLG--YKNTTISAKDILNSTLTPEFKYVNNVQRTDFVMD 211
OY 68 DLTSQDLRAATMLHTVESAFLADNLKTDIVRENTDNKLEVARLSTEGNLEDK 127
Db 212 KLSVNHRRTHLTKLMTHEVQATLRISQFQKTEFDNSTDIALKVFEPDSY-NMKIRHP 270
OY 128 PENNGHSTIQLSANTLKONGRNGEIRVAFVLYNGLPY-STFNASMKGTALSTNHS 166
Db 271 HMMN-DGDYINIPPKRAAYDSNGVAVFLYKSGIPLLSSSDNPLKPNNDNSEEE 329
OY 187 VIVNSPYITAINKFEN--KVYLADPVFTVKHIKQSEENFNPCSFWSYKRTMGY 244
Db 330 RVISS-VISVMS--SNPFLVELEKITFTLSRYVT-DRYRSLCAFWMVSPDTMGSW 384
OY 245 STGCRLLTNKTHYTTSCNHLTNFAVLA---HVEYKSDAVDLDLDVITWYGLISL 301
Db 385 SSECCELTYSNETHTSCRNHLTHFAILMSSGSPSIGIKDYN-----ILTRITOLGIITSL 439
OY 302 VCLLICIFTCFPRGLSDNRTIHKNCISLFEVALLFLGINFTDQPIACAVFAALLHF 361
Db 440 ICALICITFTWFESEIOSTRITTIHKNCISLFEVALLFLGINFTDQPIACAVFAALLHF 499
OY 362 FFLAFTWMELEGVQYIMLVEVESESRKRYFYLVGYGMPALIVASAVDRSYGTD 421
Db 500 FFLAFAFMOCIEGHLHLYVGVLYNKGFLHKNFYIFGYLSPAVVGFSAALGRRYGT 559
OY 422 KVCMLRDTYFMSFIPATLIIMLVNIFIGIALYKMFHTAILKPSGCLDNKISWVG 481
Db 560 KVCMLSTENNFMSFIPACILILVNLAFGVIIYKFRHTAGLKPEVSCFENIRSCARG 619
OY 482 AIALCLLGTWAGLWYNVESTYIMAYLFTFNSLQGMFIFPHCVLOKKVREYKCL 541
Db 620 ALALLFLGTWTWIFGVLYHVHVASVYVYALFTVSNAGFMFLFLCVLSKRIODEYTRLF 679
OY 542 R-THCCSG 548
Db 680 KNPCCCG 687

RESULT 3

```

US-09-828-366-7      : Sequence 7, Application US/09828366
:                     : Patent No. US20020010137A1
:                     :
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Goddard, Audrey
: APPLICANT: Gunney, Austin L.
: APPLICANT: Klein, Robert D.
: APPLICANT: Napier, Mary
: APPLICANT: Wood, William I.
: APPLICANT: Yvan, Jean
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC
: TITLE OF INVENTION: CELL GROWTH
: FILE REFERENCE: PI694R1C1
: CURRENT APPLICATION NUMBER: US/09/828,366
: CURRENT FILING DATE: 2001-04-05
: Prior filing data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 29
: SEQ ID NO 7
: LENGTH: 690
: TYPE: PRT
: ORGANISM: Homo Sapien
US-09-828-366-7

```

Query Match	23.9%;	Score 1088.5;	DB 10;	Length 690;
Best Local Similarity	-41.4%;	Pred. No. 1.5e-77;		
Matches 227;	Conservative 111;	Mismatches 189;	Indels 21;	Gaps 11;

QY	8	LNADITYSVAMQOLVGLDVLDRNLPRGGDSASRLNKAMVEYVNNLLQPOLNMR	67
Db	154	LSPIIDITIIYIITLAESSLLG--YKNNTISAKDITLUNSTLEEVYVNNFVORDTFVWD	211
QY	68	DLTTSOQLRAATMLHFEVESAFVLADMLKTDIYREMTDNKILKVARLSTEGNLEDKF	127
Db	212	KLAVNHRHTLTLKLMHYEQATLRISQSFQKTEBDTSTDLAKLVFFPDS--NMKHIHP	270
QY	128	PENNGHSTIOLSAKTEKONGRGEIRVAFVLYNNLGYL--STENASMKLGTEALSTNHS	186
Db	271	HMNN-DGDYIINIFPKRKAAYDSNGNVAVAFLYKYSIGELSSDNFLTKPQYDNSEEBE	329
QY	187	VIVASPTLAINKEFSN--KVYLADPVVTFYKHIKOEDEENPNPCSPFYSYKRTMTGW	244
Db	330	RVISS--VYVSMS--SNPPTLYELEKTFITFLSHKRV--DRYSJCAWPNYSPDITNGSW	384
QY	245	STOGRRLTLYTKRTTSCNHLTNPAVMA--HVEVKSADVHDLDLVLTWAGILSL	301
Db	385	SSECELTYSNEHTSCCNHLTHRALIMSSGSPSIGIDYN-----ILTRITQLGIITSL	439
QY	302	VCLLICTIFCFCEGLOSDRNTIHKNCISLFEVALLFLGINTROPACAVFALLHF	361
Db	440	ICLACITFEFMEFSEIGSTFTIHKNLCCSLFLAELVFLVGINTNTMKLFCISIIAGLHY	499
QY	362	FEIAAFITMFLGQVLYIMLYVEPESEHSRRKRYFLVYGMPALIVANSAVDYSYGD	421
Db	500	FEIAAFAMCIEGJHLIYIVGVYINKEFLKNFIFGYLSPAVVVGFSAAJGYYGTT	559
QY	422	KVCMRLDTVEYIMSEFIGPATLIMLVTLFGLATKMHHTPILKPESCGDINDKSWYG	481
Db	560	KVCMSTENNNINMSGACILILVNLALRGVYIYIKVRHRHGLKPEVSCFENINSCARG	619
QY	482	AIALLCLGLGTWAFGLAMYNESYIMAYLFTINSLQGMFLFIHFQVLQKKVREYKGL	541
Db	620	ALALLFLGTWIMIGVLHVHVASVTVALLFTVSNAFQGMFIFLFCVLSRKIQEYYRLF	679
QY	542	R-THCCSG	548
Db	680	KNVPCCF	687

Patent No. US20020132240A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gertlisen, Mary E.

APPLICANT: Goddard, A.

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth, J.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Mather, Jennie P.

APPLICANT: Pan, James

APPLICANT: Peoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/909,320

CURRENT FILING DATE: 2002-01-04

PRIOR APPLICATION NUMBER: PCT/US00/04414

PRIOR FILING DATE: 2000-02-22

PRIOR APPLICATION NUMBER: US 60/143,048

PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145,698

PRIOR FILING DATE: 1999-07-26

PRIOR APPLICATION NUMBER: US 60/146,222

PRIOR FILING DATE: 1999-07-28

PRIOR APPLICATION NUMBER: PCT/US99/20594

PRIOR FILING DATE: 1999-09-08

PRIOR APPLICATION NUMBER: PCT/US99/20944

PRIOR FILING DATE: 1999-09-13

PRIOR APPLICATION NUMBER: PCT/US99/21090

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/21547

PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/23089

PRIOR FILING DATE: 1999-10-05

PRIOR APPLICATION NUMBER: PCT/US99/28214

PRIOR FILING DATE: 1999-11-29

PRIOR APPLICATION NUMBER: PCT/US99/28313

PRIOR FILING DATE: 1999-11-30

PRIOR APPLICATION NUMBER: PCT/US99/28564

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: PCT/US99/28565

PRIOR FILING DATE: 1999-12-02

PRIOR APPLICATION NUMBER: PCT/US99/30095

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: PCT/US99/30911

PRIOR FILING DATE: 1999-12-20

PRIOR APPLICATION NUMBER: PCT/US99/30999

PRIOR FILING DATE: 1999-12-20

PRIOR APPLICATION NUMBER: PCT/US00/00219

PRIOR FILING DATE: 2000-01-05

NUMBER OF SEQ ID NOS: 423

SEQ ID NO 49

RESULT 4
US-09-909-320-49
; Sequence 49, Application US/09909320

ORGANISM: Homo sapiens
US-09-909-320-49

Db 440 ICALICITFWFSEIOSTRITTHKNCCLFLAELVFLVGINFNTNKLFCSTIINGLHY 499
Qy 362 FEIAFTWMELEGVOLYIMLVEVESEHSRRKRYLVGYPALIVANSAVDYSYSTD 421
Db 500 FEIAFAMWCEGHLHYLVGVYINKGFLHKNFYIFGLSPAIVVGSAAALGYRYGTT 559
Qy 422 KVCWLRLDTYFWSFRTGATLIMLVIFLGIALYKMFHHTAALKPESGCLDNKSWYIG 481
Db 560 KVCWLSTENNFWSFRTGACLIILVNLAFGVYIYKVRHRTAGLKPEVSCFENIRSCARG 619
Qy 482 AIALCLIGLTPWAFGLMINESTVIMAYLFTFNSLOGMFIFHCYLOKKRYKRGCL 541
Db 620 ALALFLGTWIFGVHLVHVASVYATYLVTSNAPQGMFTFLFCVLSRKIQEERYRLF 679
Qy 542 R-THCCSG 548
Db 680 KNPCCFCG 687

RESULT 6

US-09-925-300-1407
Sequence 1407, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1407
LENGTH: 713
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: SITE
LOCATION: (9)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (134)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (280)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (282)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (322)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1407

Query Match 23.9%; Score 1088.5; DB 10; Length 713;
Best Local Similarity 41.4%; Pred. No. 1.6e-77;

Matches 227; Conservative 109; Mismatches 191; Indels 21; Gaps 11;

Qy 8 LNAGDITVSAMQDLVQLDVLNLPFGKDSARSLNKAMVETVNNLLOPALNMR 67
Db 176 LSPDITIVIELAESSLLG--YKNTISAKDITLSTLFEFVTVNPFVQRODFVYWD 233
Qy 68 DLITSDDLAAATMLHIVESAFVLADNLKTDIVRENTDNKIEVARLSTEGNLEDKF 127
Db 234 KLSVNHRTHLTKIMHTEQATLRISQSFQKTTEDTNSDIALKVX-FXDSYNNKHHP 292

Qy 128 PENMGSTIOASNTLKNONGEIRVAFVLYNNIGPYL-STENASMKITGEALSTNHS 186
Db 293 HNMN-DGDYINFPKRAAYDSNGVANAFAFYTKSITGILLSSDNFLAKPOYDNDSEEE 351
Qy 187 VIVNSPVITAAINKESN--KYVLADPVYFVYKHKILOSEENPNPCSFWSYKRTMTGYW 244
Db 352 RVISS-VISVSM---SNPPLYLELEKITFTLSHRKVT-DYRSLCAFWNYSPTDMNSW 406
Qy 245 STQCGRLITTKTHTTCCSNHNTNFAVMA--HYEVKSDAVHLLDVTWVILISL 301
Db 407 SSEGCLFYYSNHTSCRNHLTHRALIMSQSPSIGIDYN----ILTRITQLOLIISL 461
Qy 302 VCLLCITFPCEGRLQSDRNTIHKNCISLPLVAELIFGINTPDQIACAVFALLHF 361
Db 462 ICALICITFWFSEIOSTRITTHKNCCLFLAELVFLVGINFNTNKLFCSTIINGLHY 521
Qy 362 FEIAFTWMELEGVOLYIMLVEVESEHSRRKRYLVGYPALIVANSAVDYSYSTD 421
Db 522 FEIAFAMWCEGHLHYLVGVYINKGFLHKNFYIFGLSPAIVVGSAAALGYRYGTT 581
Qy 422 KVCWLRLDTYFWSFRTGATLIMLVIFLGIALYKMFHHTAALKPESGCLDNKSWYIG 481
Db 582 KVCWLSTENNFWSFRTGACLIILVNLAFGVYIYKVRHRTAGLKPEVSCFENIRSCARG 641
Qy 482 AIALCLIGLTPWAFGLMINESTVIMAYLFTFNSLOGMFIFHCYLOKKRYKRGCL 541
Db 642 ALALFLGTWIFGVHLVHVASVYATYLVTSNAPQGMFTFLFCVLSRKIQEERYRLF 701
Qy 542 R-THCCSG 548
Db 702 KNPCCFCG 709

RESULT 7

US-09-978-486-5
Sequence 5, Application US/09978486
Patent No. US20020052015A1
GENERAL INFORMATION:
APPLICANT: Lin, Yi-Yun
APPLICANT: Benzer, Seymour
TITLE OF INVENTION: METHODS OF USE
FILE REFERENCE: 06618/343001
CURRENT APPLICATION NUMBER: US/09/978,486
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US/09/370,098
PRIOR FILING DATE: 1999-08-06
PRIOR APPLICATION NUMBER: 60/095,826
PRIOR FILING DATE: 1998-08-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 240
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-978-486-5

Query Match 20.3%; Score 927; DB 10; Length 240;
Best Local Similarity 67.5%; Pred. No. 1.6e-65;
Matches 162; Conservative 46; Mismatches 37; Indels 0; Gaps 0;

Qy 285 HDLIDVTWGCILSLVCLCITFCPCFRLQSDRNTIHKNCISLPLVAELFLIGIN 344
Db 1 NELLISVITWGVIVISLCLACISTFCFLRQIDRNTIHKNCINFLAELFLVIGID 60
Qy 345 RPDQIACAVFALLHFFFLAFTWMELEGVOLYIMLVEVESEHSRRKRYLVGYGPA 404
Db 61 KTOYEVACPIRGDLHYFLAFLSWCLGVALYLLVLEVESESYSKRYIYLGSCPA 120
Qy 405 LIVANSAVDYSYSTDVVCWLRLDTYFWSFRTGATLIMLVIFLGIALYKMFHHTAI 464

Db 121 LVVGIAAIDYRSYGTCKACWLRVNFYIWSIFGVPVSEVIVNLVFLMTLHKMINSSV 180
 Qy 465 LKRESCLDNIKSMVIGATALLCLGLTMAFGMYINESTIVMAYLFTTJNSLOGMEFI 524
 Db 181 LKPDSSRLDNKIKMGALGALALLGLTMAFGLLFTINKESVWVAYLFTTJNAGQVFI 240

RESULT 8 US-09-764-853-679

; Sequence 679, Application US/09764853
 ; Patent No. US20020090672A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: P1206
 ; CURRENT APPLICATION NUMBER: US/09/764,853
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 939
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 679
 ; LENGTH: 661
 ; TYPE: PRP
 ; ORGANISM: Homo sapiens
 US-09-764-853-679

Query Match 17.0%; Score 776; DB 10; Length 661;
 Best Local Similarity 33.6%; Pred. No. 5,3e-53;
 Matches 184; Conservative 102; Mismatches 229; Indels 32; Gaps 11;

Qy 35 TPGGKDSAAASLKKAVETVNNLQPOALNAMDLTTSQDLRAATMLLHVEESAFVAD 94
 Db 129 TTSSKTTGGKKELOKIVDFESILTNQTL--WTEGROEISTATYLLRDVESKVEETAL 186
 Qy 95 NLKTDIVRENTDNIKLEVARLSTEGNLEDLKPENMGHSTIQLSANLTKQNGRGEIR 154
 Db 187 KDPQGVKVLKQNSVAIEQAITDNCSEKRTNLANVQNS--MDICSDIIGDGTGQPSA 245
 Qy 155 VAEVLNINLGPYLSTENASMKLGTALSTNHSYIVNSPVITAAINKESKRYLADPVF 214
 Db 246 IATISYSLGNIT--NAF--FEEMDKKQYVYLSQVVAIGRK--RNVSLSKSVTL 297
 Qy 215 TVKHIOSEENFNPNCSFMSYSKRTMTG-YWSTQGGRLTNTKTHPTTSCNHLTNFAVLM 273
 Db 298 TFOHVKMTPESTKKVFCVYW--KSTGGSQMSRDGCFLLHVNKSHTMCSHLSFAVLM 354
 Qy 274 AHEVXHSDAVHDLDDVTIHWGILLSVCLLCIFTECFEGRLOQSDRNTIHNLCISLF 333
 Db 355 ALTSQE-----EDPVLTVITVYGLSVSLCLLLAALFLLCKAIQNTSTSLHLQSLCLF 409
 Qy 334 VAELEFLIGINTDQPIACAVFPAALLHFFFLAFTWMLFEGVOLIYI--MLVEFESEH 389
 Db 410 LAHLFLVIGIDRTEPRVLSIINGALHYLYLAFLAFTWMLLEGVHLFLARLTVVNSSIN 469
 Qy 390 SRRKY-FYLVGIGMPALIVASAAVDYRSYGTDKVCMRLDTYFIWFSIGPATLITMLNV 448
 Db 470 RLKMWIMFPGYGVPAVTAISAASMPHLGYTDRCHLHDGFMMSFLGPVCAIFSANL 529
 Qy 449 IFGLIALYKMFHHTAILKPESGCLDNIKSWIGALIALLLCLGLTMAFGMYINESTVIMA 508
 Db 530 VLEFLVFWILKRLSSLNSVSTIQTNTMLAFATQAQLFLGCTWCGILQVPAQVMA 589
 Qy 509 YLFTTJNSLOGMEFIFFHCYLOKKVREYKGCILRTHCCSGKSTESSISGSK-----TS 561
 Db 590 YLFTTJNSLOGMEFIFFHCYLOKKVREYKGCILRTHCCSGKSTESSISGSK-----TS 561
 Qy 562 GSRTPEGR 568
 Db 650 GDVFPQ 656

RESULT 9 US-09-764-898-224

; Sequence 224, Application US/09764898
 ; Patent No. US20020090673A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: P1201
 ; CURRENT APPLICATION NUMBER: US/09/764,898
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 311
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 224
 ; LENGTH: 661
 ; TYPE: PRP
 ; ORGANISM: Homo sapiens
 US-09-764-898-224

Query Match 17.0%; Score 776; DB 10; Length 661;
 Best Local Similarity 33.6%; Pred. No. 5,3e-53;
 Matches 184; Conservative 102; Mismatches 229; Indels 32; Gaps 11;

Qy 35 TPGGKDSAAASLKKAVETVNNLQPOALNAMDLTTSQDLRAATMLLHVEESAFVAD 94
 Db 129 TTSSKTTGGKKELOKIVDFESILTNQTL--WTEGROEISTATYLLRDVESKVEETAL 186
 Qy 95 NLKTDIVRENTDNIKLEVARLSTEGNLEDLKPENMGHSTIQLSANLTKQNGRGEIR 154
 Db 187 KDPQGVKVLKQNSVAIEQAITDNCSEKRTNLANVQNS--MDICSDIIGDGTGQPSA 245
 Qy 155 VAEVLNINLGPYLSTENASMKLGTALSTNHSYIVNSPVITAAINKESKRYLADPVF 214
 Db 246 IATISYSLGNIT--NAF--FEEMDKKQYVYLSQVVAIGRK--RNVSLSKSVTL 297
 Qy 215 TVKHIOSEENFNPNCSFMSYSKRTMTG-YWSTQGGRLTNTKTHPTTSCNHLTNFAVLM 273
 Db 298 TFOHVKMTPESTKKVFCVYW--KSTGGSQMSRDGCFLLHVNKSHTMCSHLSFAVLM 354
 Qy 274 AHEVXHSDAVHDLDDVTIHWGILLSVCLLCIFTECFEGRLOQSDRNTIHNLCISLF 333
 Db 355 ALTSQE-----EDPVLTVITVYGLSVSLCLLLAALFLLCKAIQNTSTSLHLQSLCLF 409
 Qy 334 VAELEFLIGINTDQPIACAVFPAALLHFFFLAFTWMLFEGVOLIYI--MLVEFESEH 389
 Db 410 LAHLFLVIGIDRTEPRVLSIINGALHYLYLAFLAFTWMLLEGVHLFLARLTVVNSSIN 469
 Qy 390 SRRKY-FYLVGIGMPALIVASAAVDYRSYGTDKVCMRLDTYFIWFSIGPATLITMLNV 448
 Db 470 RLKMWIMFPGYGVPAVTAISAASMPHLGYTDRCHLHDGFMMSFLGPVCAIFSANL 529
 Qy 449 IFGLIALYKMFHHTAILKPESGCLDNIKSWIGALIALLLCLGLTMAFGMYINESTVIMA 508
 Db 530 VLEFLVFWILKRLSSLNSVSTIQTNTMLAFATQAQLFLGCTWCGILQVPAQVMA 589
 Qy 509 YLFTTJNSLOGMEFIFFHCYLOKKVREYKGCILRTHCCSGKSTESSISGSK-----TS 561
 Db 590 YLFTTJNSLOGMEFIFFHCYLOKKVREYKGCILRTHCCSGKSTESSISGSK-----TS 561
 Qy 562 GSRTPEGR 568
 Db 650 GDVFPQ 656

RESULT 10 US-09-992-647-1

; Sequence 1, Application US/09992647
 ; Patent No. US20020146767A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Hong
 ; APPLICANT: Cohan, Victoria L.
 ; APPLICANT: Stuart, Susan G.
 ; TITLE OF INVENTION: HUMAN EMRI-LIKE G PROTEIN-COUPLED RECEPTOR
 ; FILE REFERENCE: PC-0052 CIP
 ; CURRENT APPLICATION NUMBER: US/09/992,647

```

; CURRENT FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020146767A1 429905
US-09-992-647-1

```

```

Query Match          17.0%; Score 774; DB 10; Length 652;
Best Local Similarity 33.6%; Pred. No. 7,4e-53;
Matches 184; Conservative 102; Mismatches 229; Indels 32; Gaps 11;

```

```

OY 35 TPGGDSARSINKAMVETVNNLQPOLANMRDLTTSQDLRAATMLLHTVESAFVLAD 94
DB 120 TTSSKTGGRKELQKIVDFESLITNQLT-WRTEGRQISSTATTILRDVESKYLETAL 177
OY 95 NLKTDIVRENTDNKLEVARLSTEGNLEDLKEPENMGSGTQLSANTLKGNGRGEIR 154
DB 178 KDPECKVLKIQDSVALEQATLNCSEERKTFNLVQONS-MDIRCSDIIGDTQGPVS 236
OY 155 VAFVLVNNLGPYLSTENASMKLGTEALSTNHSVIVNSPYITAINKEFENKYLADPYVE 214
DB 237 IAFIYSSLGNI---NMTF---FEEMDKKQVLYNSQVVSALGPK--RNVSLSKSVTL 288
OY 215 TVKHIKOSEENFNPCSEFWSYSKRTMTG-YWSTGCRLLTTNKTHTTSCNHLTNFAVLM 273
DB 289 TFOHKMPTSTKRVCCVW---KSTGQSGQMSRDCFLHNKSHMTMCNCHLSFAVLM 345
OY 274 AHVEKHSQAVHDLIDVTWYGIILSLVCLICFTFCEFFGLOSDRNTIHKNCISLF 333
DB 346 ALTSGE-----EDPVLYTYTVGLSVSLCLLLAATFLFLCAIDONTSTSLHQLSLCLF 400
OY 334 VAEELFLGINRTDPIACAVFAALLHFEFLAFTWMELEGVOLI---MLVEVESESH 389
DB 401 LAHLFLVIGIDRTPEKVCISITAGALHYTLAFTWMLLEGHVHLELTANLIVVNSITN 460
OY 330 SRRK-FLVLYGMPALYVSAANDYRSYGTDKVCWLRDLDTYFIWSTGPAATLIIMLV 448
DB 461 RLKMKIMRPVGVAVVVAISAASMPHLYGTADRCMLHLDGFMWSFLGPCALFSAVL 520
OY 449 IFLGALYKMFHTALILKRESCLDNIKSWVIGATALLCLLGLTNAFGMLINSEVIVA 508
DB 521 VLFILVETMLKRLSSLSNSEVSTIONTRMLAFKATRAQLFLDCTWCLGLLDVGPAAQVNA 580
OY 509 YLETFINSLOGMEFIFHCVLQKRYKKEYGKCLRTHCCSGKSTESSIGSGK-----TS 561
DB 581 YLETFINSLOGMEFIFHCVLQKRYKKEYGKCLRTHCCSGKSTESSIGSGK-----TS 561
OY 562 GSRTGGR 568
DB 641 GDFPFGQ 647

```

```

RESULT 11
US-09-992-647-11
; Sequence 11, Application US/09992647
; Patent No. US20020146767A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Hong
; APPLICANT: Cohan, Victoria L.
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: HUMAN EMRI-LIKE G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: PC-0052 CIP
; CURRENT APPLICATION NUMBER: US/09/992,647
; CURRENT FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 886

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank ID No. US20020146767A1 9784994
US-09-992-647-11

```

```

Query Match          15.7%; Score 715.5; DB 10; Length 886;
Best Local Similarity 30.4%; Pred. No. 4,6e-48;
Matches 166; Conservative 116; Mismatches 211; Indels 53; Gaps 12;

```

```

OY 47 NKAMETVNNLQPOLANMRDLTTSQDLRAATMLLHTVESAFVLAD-----NL 97
DB 366 NKTTVVSLKNTSEFVPLVKQISMWTKFKETESLATVLESVE--SMTIASFMKPSAN 423
OY 98 KTDIVRENTDNKLEVARLSTEGNLEDLKEPENMGSGTQLSANTLKGNGRGEIRVAF 157
DB 424 VTPAVR--AEYLDIESKVINKECSEENVTL-DLYAGDMKIKIGCSYIESESESTETGVAF 480
OY 158 VLYNNLGPYLSTENASMKLGTEALSTNHSVIVNSPYITAINKEFENKYL----- 207
DB 481 VSEVGMESVLN-----ERFQDH-----QAPLTSEIKKAMNSRYVGGIMTECKD 526
OY 208 -LADPVFTVKHIKOSEENFNPCSEFWSYSKRTMTGYWSTGCRLLTTNKTHTTSCNHL 266
DB 527 GFSDDPIITYLENVQPKQFERPICVSWSTDVK--GGRMITSFGVILEASEYTIICSCNQ 584
OY 267 TNPAYLMAHVEKHSQAVHDLIDVTWYGIILSLVCLICFTFCEFFGLOSDRNTIHK 326
DB 585 ANLAVTMASGEL---TYDQFSLYIISHVGIISLCLVALATATFLCHSIRHNHTYHL 639
OY 327 NCISLFLVELFLGINTDPIACAVFAALLHFEFLAFTWMELEGVOLIYIM-----L 381
DB 640 HLCVCLLAKTLPLAGIRHKTDNKTGCATITAGFLHFLACFEMLVEANILPLMVA NLKY 699
OY 382 VEFESHSRRKRYFLVGYGMPALYVSAANDYRSYGTDKVCWLRDLDTYFIWSTGPAAT 441
DB 700 VNFSSRNKIMHLICAFGLPMLVVISASVQPGYGMHNRCMLWTEGFIWISFLGYPVC 759
OY 442 LILMLNVTILGIALYKMFHTALILKRESCLDNIKSWVIGATALLCLLGLTNAFGMLYN 501
DB 760 TVIVINSLTLWTMLTILRQLSSVNAVESTLKTDLTFLTKFAPAQDLFLDCCSWGLGIFQIG 819
OY 502 ESTVIMAYLETFINSLOGMEFIFHCVLQKRYKKEYGKCL--RTHCCSGKSTESSIGSGK 559
DB 820 PAVGWAIVLETFINSLOGAFIFLHCLLNGQVREYKRNITGTKTSQSQISRIILSSM 879
OY 560 TSGSRT 565
DB 880 PSASRT 885

```

```

RESULT 12
US-09-737-149-25
; Sequence 25, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spylek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29

```

```

; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 25
; LENGTH: 3034
; TYPE: PRF
; ORGANISM: Mus musculus
; US-09-737-149-25

```

Query Match	15.4%	Score 700.5	DB 10	Length 3034
Best Local Similarity	24.5%	Pred. No. 4.4e-46		
Matches 237	Conservative 166	Mismatches 356	Indels 207	Gaps 34

QY	1	AEONTHNLNADITV--SVRADPOVLGLDLOVRLNLTGCGKDSARSLLN--KAMVETVNN	56
Db	2150	AKALRNATOGNSTLFGNDVRNRYATOLLARI--LOHESROGCFDLAATREANFHHDDVHTGSA	2208
QY	57	LLQPOLAMMDLTTSDQLRAATMLH-----TVESAFVLADNLKTDI	101
Db	2209	LLAPATEASMEQIORS--EAGAOLLRHEEAFVSNAVRNRYLTPFVLTVMILAVDI	2267
QY	102	VRE-----NNDNIKLEYAR---LSTEGULDELKPRENNGHSTIOL	139
Db	2268	FDKLMTGAQVPREDIDOEELPRELESSSPADFFKPREKK--EEDVAVLIRRTTPLTA	2326
QY	140	-----SANTLKONGRNGELTRVAF--VYNNILGPLYL---STENASMKLGEALST	183
Db	2327	QPREPARETSSSRRRRRHDERPGCAVALVYIRTLGOLLRPHYDRHDSLLRPNR---	2383
QY	184	NHSVYVNSPVTTAIAINKEFSKNVYADRVYFVKIKITGSEENFNPCSWYSTSKRT--MTG	242
Db	2384	---VINPVYVSAMVYSGTLPSPSLQRPILVEFSILETEERSKPYCVFMWNSLDTGGTG	2439
QY	243	YMSYOGCRFLTYNKNHTCTSCNHLNPNFVILMAHVEVKSDAVHDLTDVITWVGILSTV	302
Db	2440	GMSAKGCELLSRNRHYVCOGSHSASCAYLMDISRENCEV---LPKLTITTAALSLSV	2496
QY	303	CLLICITFCFFRGLDQDRNTIHKNLCTSLSYAEVLEFLTIGINTDOPRIACAVFAALLHFF	362
Db	2497	ALLVAFVLLSEVRLTSLNLSHINKMLIALFESQSLFMVGIQGTENPFLCTVVALILHYV	2556
QY	363	FLAFTWMELEGVOYLIMLYEVESEHNRKRYFLVCGMRALLYAVSAANDVRSYGTGK	422
Db	2557	SMGTFAAMLVENLHLYRMLTEVRNIDGTGMRVHVGWGPATVIGLAAGLDPOGGND	2616
QY	423	VCWMLRLDTYFIMSPFGPRLTIIMLN-VIFL---GIALKMFHNTAILKRESCLDINIKSM	478
Db	2617	FCWLSDOOLTLMSFGAPRGVYIITMIVYFVLKASAKSCQKRNHY-----YRKQVNSMLRT-	2671
QY	479	VIGAIATLLCLGLTWAFGLMYINESTVIMAYLFTFNLSQGMFIFPHCVLOKKVRYKEV-	537
Db	2672	--AFLLLLVYATWLLDLAVNOSTLTFHYHFAAFSCLOGIFVLLFHCVAAREVYKHLR	2728
QY	538	-----GKCLRTHCSCGKSPRESSIGSKTSGSRPGVYSGSOSGIRRMWMDVYRKOSSESF	593
Db	2729	AVLACKKIQLD--DSATTRATLTLTSLKCNNT--YSEBP-----DMLR-----	2767
QY	594	ITGDINSSASLNREGQLNNAADTVY---MDPLPLNGHNGNSYASIGEYLSNCOVIDR	649
Db	2768	-TALGESYASLD-----STTRREGVQKLSYSSGPARGNHGERDT-----SFIRNSMK	2814
QY	650	GYNNHETALEKKIKELTNSYIPYSLNNHENSSEONRLMAN-----KLVNVL	696
Db	2815	AHGPRSDSDSELSDHESSSYASSHTSDSEDDGGAEDKMPNAGCPAHSSTPRADALANHY	2874
QY	697	GGSGRDDAIVLDDATSFNHEESLIGEL-----IHE-----ESAPILLPVY	737
Db	2875	PAGWPDSELASDSELDTEPHLKVETVYVELHNOAQOAGNCGDRSPDESVLAKPAVY	2934

```

Oy  738  YSTEMQHPRHMYTRRIRIPDHSESEF-----PLT- -NEMTEDLOSPIRDSLY  782
Db  2935  LSSQ-----POEORRGILNNKYVYPPPLPEOPPLKSLRLREKLADCEQSTSSRT  2982
Oy  783  TSMPTLGVATESEVT- -STOREPPAKGDAEDVYKSMPLGSRNHVOLHTTYOLG  840
Db  2963  SSSLGSGGVANRDTCVITIKTPREP-----GREHLNVAMNVRTG  3022
Oy  841  RGSSDG  846
Db  3023  SAQANG  3028

```

```

RESULT 13
US-09-737-149-30
: Sequence 30. Application US/09737149
: Patent No. US20020077466A1
: GENERAL INFORMATION:
: APPLICANT: Spaderna, Steven K
: APPLICANT: Quinn, Kerry E.
: APPLICANT: Shinkets, Richard A.
: APPLICANT: Muralidhara, Padigaru
: APPLICANT: Soytek, Kimberly A.
: TITLE OR INVENTION: Polypeptides and Nucleic Acids Encoding Same
: FILE REFERENCE: 15966-620 CIP
: CURRENT APPLICATION NUMBER: US/09/737,149
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: 60/170,564
: PRIOR FILING DATE: 1999-12-14
: PRIOR APPLICATION NUMBER: 60/173,165
: PRIOR FILING DATE: 1999-12-27
: PRIOR APPLICATION NUMBER: 60/173,362
: PRIOR FILING DATE: 1999-12-27
: PRIOR APPLICATION NUMBER: 60/173,544
: PRIOR FILING DATE: 1999-12-29
: PRIOR APPLICATION NUMBER: 60/174,404
: PRIOR FILING DATE: 2000-01-04
: PRIOR APPLICATION NUMBER: 60/174,962
: PRIOR FILING DATE: 2000-01-07
: PRIOR APPLICATION NUMBER: 60/223,929
: PRIOR FILING DATE: 2000-08-09
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 30
: LENGTH: 3034
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-737-149-30

```

```

Query Match      15.4%  Score 700.5:  DB 10:  Length 3034:
Best Local Similarity 24.5%  Pred. No. 4.4e-46:
Matches 237:  Conservative 166:  Mismatches 356:  Indels 207:  Gaps

Qy      1 AECQNRNHLNAGDIIV--SVKADQVLGLDVQLRNLTPGSKDASAARSLN--KAVETVNN 56
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      2150 AKALRNNTQGSNSTLFGNDVATATVQLLARI--LQHSNQSGFDLAIRKRNPFHEDVHTGSA 2208

Qy      57 LLOQALNANMDLTTSDQLRAATMLH-----TVBSAFVLADNLKTDI 101
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      2209 LLAPATEASMEQIQRS--EAGAQLLRHFAYEFSNVARNKKTYLRPEVLTVPANMLTAVDI 2267

Qy      102 VRE-----NTDNKLEVAR--LSTEGNLEDLKFPENMGHSTIQL----- 139
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      2268 FDKLNFTGAQVPRFEDIOELPRELESSVSFPADTFKPREKK--EGPVVRLNTRTTPLLTA 2336

Qy      140 -----SANTLTQGNRNGEIVAF--VLYLNNGPYL-----STENASMKLTGEALST 183
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      2327 QPREBAERETSSSRRRRHPDEPQGFVALVALVYIRTLGQLLRPHYDPDHSLRLPNRP--- 2383

Qy      184 NHSVIVNSPYTLTAINEFNSKNYVLADPVVFYVKHIKOSEENFNPNCSFWSYSKRT--MTG 242
      || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      2384 ----VINTPVASAVNYSEGTPRLPSIQRIPLVEFSLLETLEBSKPYCVCFWMNISLDTGCTG 2439

```


Qy 747 HTTRRIRIPDHSEFPPLTNEHTEDLOSPHRDLYTSMFTLAGVAATESVTTSTQTEPP 806
Db 2897 QGSHR-----GEXPPDQESGAARLASSQPP-----PQRKGLKNNKTYTYP 2937
Qy 807 P-----AKGDAED-----VYKSPNGLSRHHVQDLHT 835
Db 2938 PLTLTEQTLKRLRELRADCEOSPSTSSLGSGPDCATVKSQREGRDLHNGVAM 2997
Qy 836 YYOLGRSSDG 846
Db 2998 NVRTGSAGADG 3008

RESULT 15
US-09-808-571A-2
; Sequence 2, Application US/09808571A
; Patent No. US20020106723A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptor for Iatrotoxin from insects
; FILE REFERENCE: Le A 34 402
; CURRENT APPLICATION NUMBER: US/09/808,571A
; CURRENT FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: DE 100 13 580.3
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-808-571A-2

Query Match 14.7%; Score 670.5; DB 10; Length 1447;
Best Local Similarity 29.2%; Pred No. 3.3e-44;
Matches 186; Conservative 121; Mismatches 268; Indels 63; Gaps 20;

Qy 39 KDSARSLNKAYETVNNLLQPOLNWMRDLTSDQLRAATMLHTVEESAFVLADNLK 98
Db 508 REAMIMELHCVKTSNLLDESQSLMDLNEDQMRVATSLTGLEVNAFLADTIIR 567
Qy 99 TDIVRENTDNKLEVARLSTEGNLEDLKPENNG---HGSTIOLSAANTLKONGRGEIRY 155
Db 568 ERSVQKVNILLSVRLVLETKTIOSSVFPDSDQWPLSDRIELPRAALIDNSEGGLVRI 627
Qy 156 AFVLYNNLGPYL--STENASMKLG--TEALSTNHS-----YIVNSPVTTAAINKE 201
Db 628 VFAAPRLLESILKPSYDHDLDKSSRYAIIISNDSDVNAIGEIOQLRLILNSKVISASLGK- 686
Qy 202 FSNKYVLADPVVFTVKHIKQSEENFNPCSFMSYSKRTMTGWSTOGCRLLTNTKTHTC 261
Db 687 -GRHIQLOSPITLTLKHLK-TEWNTNPTCVFNNYIDHA---WSANGCSLESTNRTHSVC 740
Qy 262 SCHNLTNFAVLMAHV-EVKHS--DAVHDLDDVTWVGILLSVCLLCITFPFENG--L 317
Db 741 SCHNLTNFAILMDVDEHQSLETFMDGNMRFIYISIGCVAFIYIALTLKLFNGVFV 800
Qy 318 QSDRNTIHNKLCISLFAVELLFLIGINRTDOPJACAVFALLHFFFLAFTMMLEGVOL 377
Db 801 KSARTSIYTSIYICLAIELFLIGIEGTETSLFCGFTITFLICALISGTAMFCYERFHS 860
Qy 378 YI-----MLVEFESEHSRRKYFLVGYGMPALLIVASAAVDYRSYGTDKVCM-RLDT 430
Db 861 YSTLTSDELLLEV--DQTPKVNICYLLSYGLSVAISLVIDPSTYTQNDYCVLMEANA 918
Qy 431 YFIWSTIGPATL--ITMLNVIFIGIALYKMFHHTALKPESGCLDNIKSWIGAIALLCL 488
Db 919 LFATYFVIVPLVFFVAIGYTFLSWIMCRKSRGTGLKTEHRLASVREDIRCSFVLL 978
Qy 489 LGLTAFGLMYI-----NESTVIMAYLFTFNSLOGMFFIFHCYLOKKVREYGCCLR 542
Db 979 LSAVMCSAFYLLGAKMDDOTADYGYCFICFNTLLGLTYFVHCIONEKIRREYRYR 1038

Qy 543 THC-----CSGKSTESSI--GSGRTSGSRTPGRYSTGSQSRI-----RRMNDTVRKQ 588
Db 1039 OHAMLPKCLRCRSKTSISGIVTGNPTAG--TLCVSYTSKPKPLGVSEEAHDDPQOOQ 1096
Qy 589 SESSFITGDINSASLSNREGLLNNARDTSYMDPLPLNG 626
Db 1097 QTPVPTEDAIMGATSDCE--LINEAOORRLTKSGLMTG 1132

Search completed: December 10, 2002, 11:13:51
Job time : 27 secs

Sequences producing significant alignments:

(bits) Value

gi 3882257 dbj BAA34488.1	KIAA0768 protein [Homo sapiens]	203	3e-52	L
gi 14149677 ref NP_056051.1	lectomedin-3 [Homo sapiens] >g...	202	5e-52	L
gi 7513950 pir T17199	CL3BB protein - rat >gi 3695143 gb A...	199	6e-51	
gi 7513509 pir T18398	latrophilin-3, splice variant bbag, ...	199	8e-51	
gi 7513512 pir T18408	latrophilin-3, splice variant bbbg, ...	199	8e-51	
gi 7513951 pir T17200	CL3BC protein - rat >gi 3695145 gb A...	198	1e-50	
gi 7513510 pir T18405	latrophilin-3, splice variant bbah, ...	198	1e-50	
gi 7513513 pir T18409	latrophilin-3, splice variant bbbh, ...	198	1e-50	
gi 7513947 pir T17187	CL3AB protein - rat >gi 3695137 gb A...	198	1e-50	
gi 7513506 pir T18393	latrophilin-3, splice variant abbg, ...	197	2e-50	
gi 7513503 pir T18390	latrophilin-3, splice variant abag, ...	197	2e-50	
gi 7513948 pir T17188	CL3AC protein - rat >gi 3695139 gb A...	197	2e-50	
gi 7513504 pir T18391	latrophilin-3, splice variant abah, ...	197	2e-50	
gi 7513507 pir T18394	latrophilin-3, splice variant abbh, ...	197	2e-50	
gi 18677755 ref NP_570835.1	calcium-independent alpha-latrophilin...	196	4e-50	L
gi 7513949 pir T17198	CL3BA protein - rat >gi 3695141 gb A...	196	6e-50	
gi 7513508 pir T18395	latrophilin-3, splice variant bbaf, ...	195	9e-50	
gi 7513511 pir T18407	latrophilin-3, splice variant bbbf, ...	195	1e-49	
gi 7513946 pir T17186	CL3AA protein - rat >gi 3695135 gb A...	194	1e-49	
gi 7513502 pir T18389	latrophilin-3, splice variant abaf, ...	194	2e-49	
gi 7513505 pir T18392	latrophilin-3, splice variant abbf, ...	194	2e-49	
gi 7021080 dbj BAA91375.1	unnamed protein product [Homo sa...]	150	2e-36	
gi 5880492 gb AAD54676.1 AF104938.1	lectomedin-1 beta [Homo...]	114	2e-25	L
gi 5880490 gb AAD54675.1 AF104266.1	lectomedin-1 alpha [Hom...]	114	2e-25	L
gi 6912464 ref NP_036434.1	latrophilin 1; KIAA0786 protein...	114	3e-25	L
gi 7513501 pir T18383	latrophilin-2, splice variant bbabe ...	112	8e-25	
gi 7513491 pir T18381	latrophilin-2 (splice variant bbaae)...	112	8e-25	
gi 7513498 pir T18301	latrophilin-2, splice variant baaae ...	112	9e-25	
gi 7513500 pir T18367	latrophilin-2, splice variant baabe ...	112	9e-25	
gi 7513492 pir T18382	latrophilin-2 (splice variant bbaaf)...	112	1e-24	
gi 7513493 pir T18384	latrophilin-2 (splice variant bbabf)...	112	1e-24	
gi 7513486 pir T18370	latrophilin-2 (splice variant baabf)...	112	1e-24	
gi 7513499 pir T18366	latrophilin-2, splice variant baaaf ...	112	1e-24	
gi 7513941 pir T17157	CL2AA protein - rat >gi 3695123 gb A...	106	6e-23	
gi 11280659 pir T46611	CL2BB protein - rat >gi 3695131 gb ...	106	6e-23	
gi 7513942 pir T17158	CL2AB protein - rat >gi 3695125 gb A...	106	6e-23	
gi 7513944 pir T17160	CL2BA protein - rat >gi 3695129 gb A...	106	6e-23	
gi 7513943 pir T17159	CL2AC protein - rat >gi 3695127 gb A...	106	6e-23	
gi 19705539 ref NP_599235.1	calcium-independent alpha-latrophilin...	106	7e-23	L
gi 7513945 pir T17185	CL2BC protein - rat >gi 3695133 gb A...	106	7e-23	
gi 3882293 dbj BAA34506.1	KIAA0786 protein [Homo sapiens]	103	4e-22	L
gi 7513494 pir T18385	latrophilin-2 (splice variant bbbae)...	103	5e-22	
gi 7513496 pir T18387	latrophilin-2 (splice variant bbbbe)...	103	5e-22	
gi 7513487 pir T18375	latrophilin-2 (splice variant babae)...	103	5e-22	
gi 7513489 pir T18379	latrophilin-2 (splice variant babbe)...	103	6e-22	
gi 7513495 pir T18386	latrophilin-2 (splice variant bbbaf)...	103	6e-22	
gi 7513497 pir T18388	latrophilin-2 (splice variant bbbbf)...	103	6e-22	
gi 7513488 pir T18377	latrophilin-2 (splice variant babaf)...	103	6e-22	
gi 7513490 pir T18380	latrophilin-2 (splice variant babbf)...	103	6e-22	
gi 7513938 pir T17145	CL1AB protein - rat >gi 3695117 gb A...	86	8e-17	
gi 7513940 pir T17156	CL1BB protein - rat >gi 3695121 gb A...	86	8e-17	
gi 12621148 ref NP_075251.1	CL1BA protein [Rattus norvegic...]	86	1e-16	L
gi 7513937 pir T17138	CL1AA protein - rat >gi 2213659 gb A...	86	1e-16	
gi 7662324 ref NP_055736.1	lectomedin-2; KIAA0821 protein ...	85	2e-16	L
gi 11037014 gb AAG27461.1 AF307079.1	lectomedin-2 [Homo sap...]	85	2e-16	L
gi 7513485 pir T18413	latrophilin-1, brain-specific - bovi...	84	4e-16	
gi 7513484 pir T18411	latrophilin-1, brain-specific - bovi...	84	4e-16	
gi 14043198 gb AAH07587.1 AAH07587	Unknown (protein for IMA...)	50	4e-06	L
gi 21301476 gb EAA13621.1	agCP7645 [Anopheles gambiae str....]	44	3e-04	
gi 22024081 ref NP_610397.2	CG8639 gene product [Drosophil...]	43	8e-04	L
gi 11545908 ref NP_071442.1	EGF-TM7-latrophilin-related pr...	38	0.020	L

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

<u>gi 15806511 ref NP_295222.1 </u>	NADH dehydrogenase I, G subuni...	<u>33</u>	0.81	
<u>gi 18875378 ref NP_573485.1 </u>	ETL1 [Mus musculus] >gi 144233...	<u>30</u>	8.0	L

Alignments

Get selected sequences

Select all

Deselect all

☐ >gi|3882257|dbj|BAA34488.1| KIAA0768 protein [Homo sapiens]
Length = 872

Score = 203 bits (517), Expect = 3e-52
Identities = 105/120 (87%), Positives = 105/120 (87%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLNKAMVETVXXXXXX 60
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNKAMVETV
Sbjct: 1 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTGGKDSAARSLNKAMVETVNNLLQP 60

Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
WRDLTTSND LRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG
Sbjct: 61 QALNAWRDLTTSNDQLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120

☐ >gi|14149677|ref|NP_056051.1| lectomedin-3 [Homo sapiens]
gi|11037016|gb|AAG27462.1|AF307080.1 lectomedin-3 [Homo sapiens]
Length = 1240

Score = 202 bits (515), Expect = 5e-52
Identities = 105/133 (78%), Positives = 105/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLNK----- 48
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
Sbjct: 576 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTGGKDSAARSLNKLQKRERSCRAYV 635

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTSND LRAATMLLHTVEESAFVLADNLLKTDIVRENTD
Sbjct: 636 QAMVETVNNLLQPQALNAWRDLTTSNDQLRAATMLLHTVEESAFVLADNLLKTDIVRENTD 695

Query: 108 NIKLEVARLSTEG 120
NIKLEVARLSTEG
Sbjct: 696 NIKLEVARLSTEG 708

☐ >gi|7513950|pir|T17199 CL3BB protein - rat
gi|3695143|gb|AAC62664.1| CL3BB [Rattus norvegicus]
Length = 1298

Score = 199 bits (506), Expect = 6e-51
Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLNK----- 48
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
Sbjct: 643 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTGGKDSAARSLNKLQKRERSCRAYV 702

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTSND LRAATMLL TVEESAFVLADNLLKTDIVRENTD
Sbjct: 703 QAMVETVNNLLQPQALNAWRDLTTSNDQLRAATMLLD TVEESAFVLADNLLKTDIVRENTD 762

Query: 108 NIKLEVARLSTEG 120
NI+LEVARLSTEG
Sbjct: 763 NIQLEVARLSTEG 775

☐ >gi|7513509|pir|T18398 latrophilin-3, splice variant bbag, brain-specific - bov
gi|4164067|gb|AAD05328.1| latrophilin 3 splice variant bbag [Bos taurus]

THIS PAGE BLANK (USPTO)

Length = 1299

Score = 199 bits (505), Expect = 8e-51
 Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
 Sbjct: 644 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 703

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD
 Sbjct: 704 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLDTVEESAFVLADNLLKTDIVRENTD 763

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG
 Sbjct: 764 NIQLEVARLSTEG 776

☐ >gi|7513512|pir|T18408 latrophilin-3, splice variant bbbg, brain-specific - bov
 gi|4164073|gb|AAD05331.1| latrophilin 3 splice variant bbbg [Bos taurus]
 Length = 1308

Score = 199 bits (505), Expect = 8e-51
 Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
 Sbjct: 644 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 703

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD
 Sbjct: 704 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLDTVEESAFVLADNLLKTDIVRENTD 763

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG
 Sbjct: 764 NIQLEVARLSTEG 776

☐ >gi|7513951|pir|T17200 CL3BC protein - rat
 gi|3695145|gb|AAC62665.1| CL3BC [Rattus norvegicus]
 Length = 1341

Score = 198 bits (503), Expect = 1e-50
 Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
 Sbjct: 643 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 702

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD
 Sbjct: 703 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLDTVEESAFVLADNLLKTDIVRENTD 762

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG
 Sbjct: 763 NIQLEVARLSTEG 775

☐ >gi|7513510|pir|T18405 latrophilin-3, splice variant bbah, brain-specific - bov
 gi|4164069|gb|AAD05329.1| latrophilin 3 splice variant bbah [Bos taurus]
 Length = 1342

Score = 198 bits (503), Expect = 1e-50
 Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK

THIS PAGE BLANK (USPTO)

Sbjct: 644 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRNLTTPGGKDSAARSLNKLQKRERSCRAYV 703

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD

Sbjct: 704 QAMVETVNNLLQPQALNAWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 763

Query: 108 NIKLEVARLSTEG 120
NI+LEVARLSTEG

Sbjct: 764 NIQLEVARLSTEG 776

☐ >gi|7513513|pir|T18409 latrophilin-3, splice variant bbbh, brain-specific - bov
gi|4164075|gb|AAD05332.1| latrophilin 3 splice variant bbbh [Bos taurus]
Length = 1351

Score = 198 bits (503), Expect = 1e-50
Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK

Sbjct: 644 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRNLTTPGGKDSAARSLNKLQKRERSCRAYV 703

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD

Sbjct: 704 QAMVETVNNLLQPQALNAWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 763

Query: 108 NIKLEVARLSTEG 120
NI+LEVARLSTEG

Sbjct: 764 NIQLEVARLSTEG 776

☐ >gi|7513947|pir|T17187 CL3AB protein - rat
gi|3695137|gb|AAC62661.1| CL3AB [Rattus norvegicus]
Length = 1230

Score = 198 bits (503), Expect = 1e-50
Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK

Sbjct: 575 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRNLTTPGGKDSAARSLNKLQKRERSCRAYV 634

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD

Sbjct: 635 QAMVETVNNLLQPQALNAWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 694

Query: 108 NIKLEVARLSTEG 120
NI+LEVARLSTEG

Sbjct: 695 NIQLEVARLSTEG 707

☐ >gi|7513506|pir|T18393 latrophilin-3, splice variant abbg, brain-specific - bov
gi|4164061|gb|AAD05325.1| latrophilin 3 splice variant abbg [Bos taurus]
Length = 1240

Score = 197 bits (502), Expect = 2e-50
Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK

Sbjct: 576 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRNLTTPGGKDSAARSLNKLQKRERSCRAYV 635

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD

Sbjct: 636 QAMVETVNNLLQPQALNAWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 695

Query: 108 NIKLEVARLSTEG 120

THIS PAGE BLANK (USPTO)

NI+LEVARLSTEG

Sbjct: 696 NIQLEVARLSTEG 708

☐ >gi|7513503|pir||T18390 latrophilin-3, splice variant abag, brain-specific - bov
 gi|4164055|gb|AAD05322.1| latrophilin 3 splice variant abag [Bos taurus]
 Length = 1231

Score = 197 bits (502), Expect = 2e-50

Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
 Sbjct: 576 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 635

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD
 Sbjct: 636 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLD TVEESAFVLADNLLKTDIVRENTD 695

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG
 Sbjct: 696 NIQLEVARLSTEG 708

☐ >gi|7513948|pir||T17188 CL3AC protein - rat
 gi|3695139|gb|AAC62662.1| CL3AC [Rattus norvegicus]
 Length = 1273

Score = 197 bits (501), Expect = 2e-50

Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
 Sbjct: 575 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 634

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD
 Sbjct: 635 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLD TVEESAFVLADNLLKTDIVRENTD 694

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG
 Sbjct: 695 NIQLEVARLSTEG 707

☐ >gi|7513504|pir||T18391 latrophilin-3, splice variant abah, brain-specific - bov
 gi|4164057|gb|AAD05323.1| latrophilin 3 splice variant abah [Bos taurus]
 Length = 1274

Score = 197 bits (501), Expect = 2e-50

Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
 Sbjct: 576 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 635

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD
 Sbjct: 636 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLD TVEESAFVLADNLLKTDIVRENTD 695

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG
 Sbjct: 696 NIQLEVARLSTEG 708

☐ >gi|7513507|pir||T18394 latrophilin-3, splice variant abbh, brain-specific - bov
 gi|4164063|gb|AAD05326.1| latrophilin 3 splice variant abbh [Bos taurus]
 Length = 1283

THIS PAGE BLANK (USPTO)

Score = 197 bits (501), Expect = 2e-50
Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
Sbjct: 576 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 635

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTS LRAATMLL TVEESAFVLADNLLKTDIVRENTD
Sbjct: 636 QAMVETVNNLLQPQALNAWRDLTTSQDLRAATMLLD TVEESAFVLADNLLKTDIVRENTD 695

Query: 108 NIKLEVARLSTEG 120
NI+LEVARLSTEG
Sbjct: 696 NIQLEVARLSTEG 708

☐ >gi|18677755|ref|NP_570835.1| calcium-independent alpha-latrotoxin receptor hom
lectomedin-3 [Rattus norvegicus]
gi|7513906|pir|T14327 alpha-latrotoxin receptor 3, calcium-independent - rat
gi|3882981|gb|AAC77816.1| calcium-independent alpha-latrotoxin receptor homolog
norvegicus]
Length = 1550

Score = 196 bits (499), Expect = 4e-50
Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
Sbjct: 643 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 702

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTS LRAATMLL TVEESAFVLADNLLKTDIVRENTD
Sbjct: 703 QAMVETVNNLLQPQALNAWRDLTTSQDLRAATMLLD TVEESAFVLADNLLKTDIVRENTD 762

Query: 108 NIKLEVARLSTEG 120
NI+LEVARLSTEG
Sbjct: 763 NIQLEVARLSTEG 775

☐ >gi|7513949|pir|T17198 CL3BA protein - rat
gi|3695141|gb|AAC62663.1| CL3BA [Rattus norvegicus]
Length = 1527

Score = 196 bits (497), Expect = 6e-50
Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
Sbjct: 643 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 702

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTD 107
AMVETV WRDLTTS LRAATMLL TVEESAFVLADNLLKTDIVRENTD
Sbjct: 703 QAMVETVNNLLQPQALNAWRDLTTSQDLRAATMLLD TVEESAFVLADNLLKTDIVRENTD 762

Query: 108 NIKLEVARLSTEG 120
NI+LEVARLSTEG
Sbjct: 763 NIQLEVARLSTEG 775

☐ >gi|7513508|pir|T18395 latrophilin-3, splice variant bbaf, brain-specific - bov
gi|4164065|gb|AAD05327.1| latrophilin 3 splice variant bbaf [Bos taurus]
Length = 1571

Score = 195 bits (496), Expect = 9e-50
Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48

THIS PAGE BLANK (USPTO)

AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
 Sbjct: 644 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRNLTTPGGKDSAARSLNKLQKRERSCRAYV 703

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD

Sbjct: 704 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLDTVEESAFVLADNLLKTDIVRENTD 763

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG

Sbjct: 764 NIQLEVARLSTEG 776

☐ >gi|7513511|pir|T18407 latrophilin-3, splice variant bbbf, brain-specific - bov
 gi|4164071|gb|AAD05330.1| latrophilin 3 splice variant bbbf [Bos taurus]
 Length = 1580

Score = 195 bits (495), Expect = 1e-49
 Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLNRNLTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK

Sbjct: 644 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRNLTTPGGKDSAARSLNKLQKRERSCRAYV 703

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD

Sbjct: 704 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLDTVEESAFVLADNLLKTDIVRENTD 763

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG

Sbjct: 764 NIQLEVARLSTEG 776

☐ >gi|7513946|pir|T17186 CL3AA protein - rat
 gi|3695135|gb|AAC62660.1| CL3AA [Rattus norvegicus]
 Length = 1459

Score = 194 bits (494), Expect = 1e-49
 Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLNRNLTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK

Sbjct: 575 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRNLTTPGGKDSAARSLNKLQKRERSCRAYV 634

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD

Sbjct: 635 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLDTVEESAFVLADNLLKTDIVRENTD 694

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG

Sbjct: 695 NIQLEVARLSTEG 707

☐ >gi|7513502|pir|T18389 latrophilin-3, splice variant abaf, brain-specific - bov
 gi|4164053|gb|AAD05321.1| latrophilin 3 splice variant abaf [Bos taurus]
 Length = 1503

Score = 194 bits (493), Expect = 2e-49
 Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLNRNLTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK

Sbjct: 576 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRNLTTPGGKDSAARSLNKLQKRERSCRAYV 635

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMMLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD

Sbjct: 636 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLDTVEESAFVLADNLLKTDIVRENTD 695

THIS PAGE BLANK (USPTO)

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG
 Sbjct: 696 NIQLEVARLSTEG 708

☐ >gi|7513505|pir|T18392 latrophilin-3, splice variant abbf, brain-specific - bov
 gi|4164059|gb|AAD05324.1| latrophilin 3 splice variant abbf [Bos taurus]
 Length = 1512

Score = 194 bits (493), Expect = 2e-49
 Identities = 103/133 (77%), Positives = 104/133 (78%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNK----- 48
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNK
 Sbjct: 576 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKLQKRERSCRAYV 635

Query: 49 -AMVETVXXXXXXXXXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 AMVETV WRDLTTSN LRAATMLL TVEESAFVLADNLLKTDIVRENTD
 Sbjct: 636 QAMVETVNNLLQPQALNAWRDLTTSNQLRAATMLLD TVEESAFVLADNLLKTDIVRENTD 695

Query: 108 NIKLEVARLSTEG 120
 NI+LEVARLSTEG
 Sbjct: 696 NIQLEVARLSTEG 708

☐ >gi|7021080|dbj|BAA91375.1| unnamed protein product [Homo sapiens]
 Length = 393

Score = 150 bits (380), Expect = 2e-36
 Identities = 77/92 (83%), Positives = 77/92 (83%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNKAMVETVXXXXXX 60
 AE TRNHLNAGDITYSVRAMD LVGLLDV LRNLTPGGKDSAARSLNKAMVETV
 Sbjct: 262 AEQTRNHLNAGDITYSVRAMDQLVGLLDVQLRLNLTTPGGKDSAARSLNKAMVETVNNLLQP 321

Query: 61 XXXXXWRDLTTSNLRATMLLHTVEESAFVL 92
 WRDLTTSN LRAATMLLHTVEESAFVL
 Sbjct: 322 QALNAWRDLTTSNQLRAATMLLHTVEESAFVL 353

☐ >gi|5880492|gb|AAD54676.1|AF104938.1 lectomedin-1 beta [Homo sapiens]
 Length = 1123

Score = 114 bits (286), Expect = 2e-25
 Identities = 60/120 (50%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNKAMVETVXXXXXX 60
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
 Sbjct: 547 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVD TVDNLLRP 606

Query: 61 XXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
 W+ + +S+ ATMLL T+EE AFVLADNLL+ V T+NI LEVA LSTEG
 Sbjct: 607 EALESWKHMNSSEQAHTATMLLD TLEEGAFVLADNLLLEP TRVSMPTENIVLEVA VLSTEG 666

☐ >gi|5880490|gb|AAD54675.1|AF104266.1 lectomedin-1 alpha [Homo sapiens]
 Length = 1177

Score = 114 bits (286), Expect = 2e-25
 Identities = 60/120 (50%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNKAMVETVXXXXXX 60
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
 Sbjct: 547 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVD TVDNLLRP 606

Query: 61 XXXXXWRDLTTSNLRATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
 W+ + +S+ ATMLL T+EE AFVLADNLL+ V T+NI LEVA LSTEG
 Sbjct: 607 EALESWKHMNSSEQAHTATMLLD TLEEGAFVLADNLLLEP TRVSMPTENIVLEVA VLSTEG 666

THIS PAGE BLANK (USPTO)

☐ >gi|6912464|ref|NP_036434.1| latrophilin 1; KIAA0786 protein; lectomedin-1; lat sapiens]
 gi|4034486|emb|CAA10458.1| latrophilin-2 [Homo sapiens]
 gi|5880494|gb|AAD54677.1|AF104939.1 lectomedin-1 gamma [Homo sapiens]
 gi|6274511|emb|CAB60229.1| latrophilin-2 [Homo sapiens]
 Length = 1403

Score = 114 bits (284), Expect = 3e-25
 Identities = 60/120 (50%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLNKAMVETVXXXXXX 60
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
 Sbjct: 547 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVDNLLRP 606
 Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
 W+ + +S+ ATMLL T+EE AFVLADNLL+ V T+NI LEVA LSTEG
 Sbjct: 607 EALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLPEPTRVSMPTENIVLEVAVLSTEG 666

☐ >gi|7513501|pir|T18383 latrophilin-2, splice variant bbabe - bovine
 gi|4164041|gb|AAD05315.1| Latrophilin 2 splice variant bbabe [Bos taurus]
 Length = 1422

Score = 112 bits (280), Expect = 8e-25
 Identities = 59/120 (49%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLNKAMVETVXXXXXX 60
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
 Sbjct: 551 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVDNLLRP 610
 Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
 W+ + +S+ ATMLL T+EE AFVLADNLL+ V T+NI LEVA LSTEG
 Sbjct: 611 EALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLPEPTRVSMPTENIVLEVAVLSTEG 670

☐ >gi|7513491|pir|T18381 latrophilin-2 (splice variant bbaae) - bovine
 gi|4164037|gb|AAD05313.1| latrophilin 2 splice variant bbaae [Bos taurus]
 Length = 1407

Score = 112 bits (280), Expect = 8e-25
 Identities = 59/120 (49%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLNKAMVETVXXXXXX 60
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
 Sbjct: 551 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVDNLLRP 610
 Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
 W+ + +S+ ATMLL T+EE AFVLADNLL+ V T+NI LEVA LSTEG
 Sbjct: 611 EALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLPEPTRVSMPTENIVLEVAVLSTEG 670

☐ >gi|7513498|pir|T18301 latrophilin-2, splice variant baaae - bovine
 gi|4164021|gb|AAD05305.1| latrophilin 2 splice variant baaae [Bos taurus]
 Length = 1341

Score = 112 bits (280), Expect = 9e-25
 Identities = 59/120 (49%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLNKAMVETVXXXXXX 60
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
 Sbjct: 485 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVDNLLRP 544
 Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
 W+ + +S+ ATMLL T+EE AFVLADNLL+ V T+NI LEVA LSTEG
 Sbjct: 545 EALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLPEPTRVSMPTENIVLEVAVLSTEG 604

THIS PAGE BLANK (USPTO)

☐ >gi|7513500|pir|T18367 latrophilin-2, splice variant baabe - bovine
gi|4164025|gb|AAD05307.1| latrophilin 2 splice variant baabe [Bos taurus]
Length = 1356

Score = 112 bits (280), Expect = 9e-25
Identities = 59/120 (49%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNKAMVETVXXXXXX 60
A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
Sbjct: 485 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVNDLLRP 544
Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
W+ + +S+ ATMLL T+EE AFVLADNL++ V T+NI LEVA LSTEG
Sbjct: 545 EALESWKHMNSSEQAHTATMLLDLTLEEGAFVLADNLVEPTRVSMPTENIVLEVAVLSTEG 604

☐ >gi|7513492|pir|T18382 latrophilin-2 (splice variant bbaaf) - bovine
gi|4164039|gb|AAD05314.1| latrophilin 2 splice variant bbaaf [Bos taurus]
Length = 1450

Score = 112 bits (280), Expect = 1e-24
Identities = 59/120 (49%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNKAMVETVXXXXXX 60
A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
Sbjct: 551 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVNDLLRP 610
Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
W+ + +S+ ATMLL T+EE AFVLADNL++ V T+NI LEVA LSTEG
Sbjct: 611 EALESWKHMNSSEQAHTATMLLDLTLEEGAFVLADNLVEPTRVSMPTENIVLEVAVLSTEG 670

☐ >gi|7513493|pir|T18384 latrophilin-2 (splice variant bbabf) - bovine
gi|4164043|gb|AAD05316.1| latrophilin 2 splice variant bbabf [Bos taurus]
Length = 1465

Score = 112 bits (280), Expect = 1e-24
Identities = 59/120 (49%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNKAMVETVXXXXXX 60
A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
Sbjct: 551 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVNDLLRP 610
Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
W+ + +S+ ATMLL T+EE AFVLADNL++ V T+NI LEVA LSTEG
Sbjct: 611 EALESWKHMNSSEQAHTATMLLDLTLEEGAFVLADNLVEPTRVSMPTENIVLEVAVLSTEG 670

☐ >gi|7513486|pir|T18370 latrophilin-2 (splice variant baabf) - bovine
gi|4164027|gb|AAD05308.1| latrophilin 2 splice variant baabf [Bos taurus]
Length = 1399

Score = 112 bits (280), Expect = 1e-24
Identities = 59/120 (49%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNKAMVETVXXXXXX 60
A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
Sbjct: 485 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVNDLLRP 544
Query: 61 XXXXXWRDLTTSNDLRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
W+ + +S+ ATMLL T+EE AFVLADNL++ V T+NI LEVA LSTEG
Sbjct: 545 EALESWKHMNSSEQAHTATMLLDLTLEEGAFVLADNLVEPTRVSMPTENIVLEVAVLSTEG 604

☐ >gi|7513499|pir|T18366 latrophilin-2, splice variant baaaf - bovine
gi|4164023|gb|AAD05306.1| latrophilin 2 splice variant baaaf [Bos taurus]

THIS PAGE BLANK (USPTO)

Length = 1384

Score = 112 bits (280), Expect = 1e-24
 Identities = 59/120 (49%), Positives = 78/120 (65%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLNKAMVETVXXXXXX 60
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS NKA+V+TV
 Sbjct: 485 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKAIVDTVDNLLRP 544

Query: 61 XXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTDNIKLEVARLSTEG 120
 W+ + +S+ ATMLL T+EE AFVLADNLL+ V T+NI LEVA LSTEG
 Sbjct: 545 EALESWKHMNSSEQAHTATMLLDLTLEEGAFVLADNLEPTRVSMPTENIVLEVAVLSTEG 604

☐ >gi|7513941|pir|T17157 CL2AA protein - rat
 gi|3695123|gb|AAC62654.1| CL2AA [Rattus norvegicus]
 Length = 1452

Score = 106 bits (265), Expect = 6e-23
 Identities = 61/133 (45%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLN----- 47
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
 Sbjct: 551 AKHTKGTVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNLL+ V TD
 Sbjct: 611 KAIVDTVDNLLRAETLDCWKHMNSSEQAHTATMLLDLTLEEGAFVLADNLEPTRVSMPTD 670

Query: 108 NIKLEVARLSTEG 120
 NI LEVA LSTEG
 Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|11280659|pir|T46611 CL2BB protein - rat
 gi|3695131|gb|AAC62658.1| CL2BB [Rattus norvegicus]
 Length = 1435

Score = 106 bits (264), Expect = 6e-23
 Identities = 61/133 (45%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLN----- 47
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
 Sbjct: 551 AKHTKGTVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNLL+ V TD
 Sbjct: 611 KAIVDTVDNLLRAETLDCWKHMNSSEQAHTATMLLDLTLEEGAFVLADNLEPTRVSMPTD 670

Query: 108 NIKLEVARLSTEG 120
 NI LEVA LSTEG
 Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|7513942|pir|T17158 CL2AB protein - rat
 gi|3695125|gb|AAC62655.1| CL2AB [Rattus norvegicus]
 Length = 1420

Score = 106 bits (264), Expect = 6e-23
 Identities = 61/133 (45%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLN----- 47
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
 Sbjct: 551 AKHTKGTVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNLL+ V TD

THIS PAGE BLANK (USPTO)

Sbjct: 611 KAIVDTVDNLLRAETLDCWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLLEPTRVSMPTD 670

Query: 108 NIKLEVARLSTEG 120

NI LEVA LSTEG

Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|7513944|pir|T17160 CL2BA protein - rat
 gi|3695129|gb|AAC62657.1| CL2BA [Rattus norvegicus]
 Length = 1467

Score = 106 bits (264), Expect = 6e-23

Identities = 61/133 (45%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNTPGGKDSAARSLN----- 47

A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N

Sbjct: 551 AKHTKGTVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107

KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNLL+ V TD

Sbjct: 611 KAIVDTVDNLLRAETLDCWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLLEPTRVSMPTD 670

Query: 108 NIKLEVARLSTEG 120

NI LEVA LSTEG

Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|7513943|pir|T17159 CL2AC protein - rat
 gi|3695127|gb|AAC62656.1| CL2AC [Rattus norvegicus]
 Length = 1463

Score = 106 bits (264), Expect = 6e-23

Identities = 61/133 (45%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNTPGGKDSAARSLN----- 47

A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N

Sbjct: 551 AKHTKGTVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107

KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNLL+ V TD

Sbjct: 611 KAIVDTVDNLLRAETLDCWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLLEPTRVSMPTD 670

Query: 108 NIKLEVARLSTEG 120

NI LEVA LSTEG

Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|19705539|ref|NP_599235.1| calcium-independent alpha-latrotoxin receptor hom
 norvegicus]
 gi|7513907|pir|T14324 alpha-latrotoxin receptor, calcium-independent - rat
 gi|3766205|gb|AAC77815.1| calcium-independent alpha-latrotoxin receptor homolog
 norvegicus]
 Length = 1487

Score = 106 bits (264), Expect = 7e-23

Identities = 61/133 (45%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNTPGGKDSAARSLN----- 47

A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N

Sbjct: 551 AKHTKGTVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107

KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNLL+ V TD

Sbjct: 611 KAIVDTVDNLLRAETLDCWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLLEPTRVSMPTD 670

Query: 108 NIKLEVARLSTEG 120

NI LEVA LSTEG

THIS PAGE BLANK (USPTO)

Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|7513945|pir|T17185 CL2BC protein - rat
gi|3695133|gb|AAC62659.1| CL2BC [Rattus norvegicus]
Length = 1478

Score = 106 bits (264), Expect = 7e-23
Identities = 61/133 (45%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLN----- 47
A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
Sbjct: 551 AKHTKGTVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610
Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNLL+ V TD
Sbjct: 611 KAIVDTVDNLLRAETLDCWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLLEPTRVSMPTD 670
Query: 108 NIKLEVARLSTEG 120
NI LEVA LSTEG
Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|3882293|dbj|BAA34506.1| KIAA0786 protein [Homo sapiens]
Length = 1021

Score = 103 bits (257), Expect = 4e-22
Identities = 60/133 (45%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLN----- 47
A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
Sbjct: 109 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 168
Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNLL+ V T+
Sbjct: 169 KAIVDTVDNLLRPEALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLLLEPTRVSMPTD 228
Query: 108 NIKLEVARLSTEG 120
NI LEVA LSTEG
Sbjct: 229 NIVLEVAVLSTEG 241

☐ >gi|7513494|pir|T18385 latrophilin-2 (splice variant bbbae) - bovine
gi|4164045|gb|AAD05317.1| latrophilin 2 splice variant bbbae [Bos taurus]
Length = 1420

Score = 103 bits (256), Expect = 5e-22
Identities = 59/133 (44%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTGGKDSAARSLN----- 47
A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
Sbjct: 551 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610
Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNL++ V T+
Sbjct: 611 KAIVDTVDNLLRPEALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLVEPTRVSMPTD 670
Query: 108 NIKLEVARLSTEG 120
NI LEVA LSTEG
Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|7513496|pir|T18387 latrophilin-2 (splice variant bbbbe) - bovine
gi|4164049|gb|AAD05319.1| latrophilin 2 splice variant bbbbe [Bos taurus]
Length = 1435

Score = 103 bits (256), Expect = 5e-22
Identities = 59/133 (44%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

THIS PAGE BLANK (USPTO)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLN----- 47
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
 Sbjct: 551 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNL++ V T+
 Sbjct: 611 KAIVDTVDNLLRPEALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLVEPTRVSMPT 670

Query: 108 NIKLEVARLSTEG 120
 NI LEVA LSTEG
 Sbjct: 671 NIVLEVAVLSTEG 683

☐ >gi|7513487|pir|T18375 latrophilin-2 (splice variant babae) - bovine
 gi|4164029|gb|AAD05309.1| latrophilin 2 splice variant babae [Bos taurus]
 Length = 1354

Score = 103 bits (256), Expect = 5e-22
 Identities = 59/133 (44%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLN----- 47
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
 Sbjct: 485 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 544

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNL++ V T+
 Sbjct: 545 KAIVDTVDNLLRPEALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLVEPTRVSMPT 604

Query: 108 NIKLEVARLSTEG 120
 NI LEVA LSTEG
 Sbjct: 605 NIVLEVAVLSTEG 617

☐ >gi|7513489|pir|T18379 latrophilin-2 (splice variant babbe) - bovine
 gi|4164033|gb|AAD05311.1| latrophilin 2 splice variant babbe [Bos taurus]
 Length = 1369

Score = 103 bits (256), Expect = 6e-22
 Identities = 59/133 (44%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLN----- 47
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
 Sbjct: 485 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 544

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNL++ V T+
 Sbjct: 545 KAIVDTVDNLLRPEALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLVEPTRVSMPT 604

Query: 108 NIKLEVARLSTEG 120
 NI LEVA LSTEG
 Sbjct: 605 NIVLEVAVLSTEG 617

☐ >gi|7513495|pir|T18386 latrophilin-2 (splice variant bbbaf) - bovine
 gi|4164047|gb|AAD05318.1| latrophilin 2 splice variant bbbaf [Bos taurus]
 Length = 1463

Score = 103 bits (256), Expect = 6e-22
 Identities = 59/133 (44%), Positives = 78/133 (58%), Gaps = 13/133 (9%)

Query: 1 AENTRNHLNAGDITYSVRAMDNLVGLLDVNLRLNLTTPGGKDSAARSLN----- 47
 A++T+ + AGD++ SVR M+ LV +LD L+ L P KDSA RS N
 Sbjct: 551 AKHTKGPVFAGDVSSSVRLMEQLVDILDAQLQELKPSEKDSAGRSYNKLQKREKTCRAYL 610

Query: 48 KAMVETVXXXXXXXXXXXXWRDLTTSNLRRAATMLLHTVEESAFVLADNLLKTDIVRENTD 107
 KA+V+TV W+ + +S+ ATMLL T+EE AFVLADNL++ V T+
 Sbjct: 611 KAIVDTVDNLLRPEALESWKHMNSSEQAHTATMLLDTLEEGAFVLADNLVEPTRVSMPT 670

THIS PAGE BLANK (USPTO)